Le matematiche

Contenuti, metodi e significato

A cura di
A. D. Aleksandrov · A. N. Kolmogorov
M. A. Lavrent’ev

Prefazione e capitoli I–XX

Traduzione italiana di lavoro

Con indice analitico e dei nomi

Nota a questa versione

Questo PDF riunisce la prefazione all'edizione russa e i primi venti capitoli: «Uno sguardo generale sulla matematica», di A. D. Aleksandrov, «Analisi», di M. A. Lavrent'ev e S. M. Nikol'skij, «Geometria analitica» e «Algebra: teoria delle equazioni algebriche», di B. N. Delone, ed «Equazioni differenziali ordinarie», di I. G. Petrovskij, ed «Equazioni alle derivate parziali», di S. L. Sobolev e O. A. Ladyženskaja, «Curve e superfici», di A. D. Aleksandrov, «Il calcolo delle variazioni», di V. I. Krylov, «Funzioni di variabile complessa», di M. V. Keldyš, «I numeri primi», di K. K. Mardžanišvili e A. B. Postnikov, «La teoria della probabilità», di A. N. Kolmogorov, «L'approssimazione delle funzioni», di S. M. Nikol'skij, «Metodi di approssimazione e tecniche di calcolo», di V. I. Krylov, «Le macchine calcolatrici elettroniche», di S. A. Lebedev e L. V. Kantorovič, «Teoria delle funzioni di una variabile reale», di S. B. Stečkin, «Algebra lineare», di D. K. Faddeev, «Geometria non euclidea», di A. D. Aleksandrov, «Topologia», di P. S. Aleksandrov, «Analisi funzionale», di I. M. Gel'fand, e «Gruppi e altri sistemi algebrici», di A. I. Mal'cev.

Il testo è tradotto dalla versione inglese del libro fornito. La copia DjVu è stata affiancata dalla corrispondente edizione elettronica per recuperare testo, formule e illustrazioni. La traduzione è stata preparata con assistenza automatica e successiva revisione linguistica e terminologica.

Sono conservate le sezioni dei venti capitoli, le formule, le figure, le tavole, le note e le bibliografie. Le correzioni di alcuni refusi matematici dei capitoli II–XX sono dichiarate nelle rispettive note editoriali. Le note sono raccolte alla fine del capitolo e raggiungibili dai richiami nel testo. La numerazione delle pagine appartiene a questa versione italiana.

Le valutazioni storiche e i riferimenti al presente rispecchiano l'epoca del testo originale. Espressioni come «il secolo scorso» e «il nostro secolo» indicano rispettivamente l'Ottocento e il Novecento. I riferimenti ai capitoli successivi restano quelli dell'opera completa.

I titoli delle opere nella bibliografia sono mantenuti nella lingua delle edizioni citate. Le osservazioni sulle scelte della traduzione inglese sono esplicitamente attribuite a quell'edizione.

Indice

Prefazione all'edizione russa 12

Uno sguardo generale sulla matematica 14

§ 1. I caratteri distintivi della matematica . . . 14

§ 2. L'aritmetica . . . 20

§ 3. La geometria . . . 31

§ 4. Aritmetica e geometria . . . 35

§ 5. L'epoca della matematica elementare . . . 45

§ 6. La matematica delle grandezze variabili . . . 52

§ 7. La matematica contemporanea . . . 65

Letture consigliate . . . 73

Note al capitolo I . . . 74

Analisi 77

§ 1. Introduzione . . . 77

§ 2. Funzione . . . 84

§ 3. Limiti . . . 92

§ 4. Funzioni continue . . . 99

§ 5. Derivata . . . 104

§ 6. Regole di derivazione . . . 113

§ 7. Massimi e minimi; studio dei grafici delle funzioni . . . 120

§ 8. Incremento e differenziale di una funzione 129

§ 9. Formula di Taylor 136

§ 10. Integrale 141

§ 11. Integrali indefiniti; tecniche di integrazione 149

§ 12. Funzioni di più variabili 154

§ 13. Generalizzazioni del concetto di integrale 169

§ 14. Serie 177

Letture consigliate 191

Note al capitolo II 192

Geometria analitica 196

§ 1. Introduzione 196

§ 2. Le due idee fondamentali di Cartesio 197

§ 3. Problemi elementari 199

§ 4. Studio delle curve rappresentate da equazioni di primo e secondo grado 201

§ 5. Il metodo di Cartesio per risolvere le equazioni algebriche di terzo e quarto grado 204

§ 6. La teoria generale dei diametri di Newton 207

§ 7. Ellisse, iperbole e parabola 210

§ 8. Riduzione dell'equazione generale di secondo grado alla forma canonica 228

§ 9. Rappresentazione di forze, velocità e accelerazioni mediante terne di numeri; teoria dei vettori 234

§ 10. Geometria analitica nello spazio; equazioni di superfici e curve 241

§ 11. Trasformazioni affini e ortogonali 253

§ 12. Teoria degli invarianti 265

§ 13. Geometria proiettiva 270

§ 14. Trasformazioni di Lorentz 277

Conclusione 286

Letture consigliate 289

Note al capitolo III 290

Algebra: teoria delle equazioni algebriche 291

§ 1. Introduzione 291

§ 2. Risoluzione algebrica di un'equazione 295

§ 3. Il teorema fondamentale dell'algebra 309

§ 4. Studio della distribuzione delle radici di un polinomio nel piano complesso 321

§ 5. Calcolo approssimato delle radici 332

Letture consigliate 339

Note al capitolo IV 340

Equazioni differenziali ordinarie 342

§ 1. Introduzione 342

§ 2. Equazioni differenziali lineari a coefficienti costanti 353

§ 3. Osservazioni generali sulla formulazione e sulla risoluzione delle equazioni differenziali . . 359

§ 4. Interpretazione geometrica del problema di integrazione delle equazioni differenziali e sua generalizzazione 361

§ 5. Esistenza e unicità della soluzione di un'equazione differenziale; risoluzione approssimata . 365

§ 6. Punti singolari 371

§ 7. Teoria qualitativa delle equazioni differenziali ordinarie 377

Letture consigliate 384

Note al capitolo V 385

EQUAZIONI ALLE DERIVATE PARZIALI 386

§1. Introduzione 386

§2. Le equazioni più semplici della fisica matematica 388

§3. Problemi ai valori iniziali e al contorno; unicità della soluzione 397

§4. La propagazione delle onde 406

§5. Metodi di costruzione delle soluzioni 408

§6. Soluzioni generalizzate 427

Letture consigliate 433

Note al capitolo VI 434

Curve e superfici 436

§1. Oggetti e metodi della teoria delle curve e delle superfici 436

§2. La teoria delle curve 440

§3. Concetti fondamentali della teoria delle superfici 453

§4. Geometria intrinseca e deformazione delle superfici 468

§5. Nuovi sviluppi della teoria delle curve e delle superfici 483

Letture consigliate 490

Note al capitolo VII 491

Il calcolo delle variazioni 494

§1. Introduzione 494

§2. Le equazioni differenziali del calcolo delle variazioni 499

§3. Metodi di soluzione approssimata dei problemi del calcolo delle variazioni 511

Letture consigliate 513

Note al capitolo VIII 514

Funzioni di variabile complessa 515

§1. Numeri complessi e funzioni di variabile complessa 515

§2. Le funzioni complesse e i problemi della fisica matematica 528

§3. Le funzioni complesse e la geometria 539

§4. L'integrale curvilineo; la formula di Cauchy e le sue conseguenze 549

§5. Unicità e prolungamento analitico 561

§6. Conclusione 567

Letture consigliate 567

Note al capitolo IX 568

I numeri primi 570

§1. Lo studio della teoria dei numeri 570

§2. Le ricerche sui numeri primi 574

§3. Il metodo di Čebyšev 581

§4. Il metodo di Vinogradov 587

§5. Somme di due quadrati e interi complessi 594

Letture consigliate 596

Note al capitolo X 597

La teoria della probabilità 599

§1. Le leggi della probabilità 599

§2. Assiomi e formule fondamentali della teoria elementare della probabilità 601

§3. La legge dei grandi numeri e i teoremi limite 607

§4. Ulteriori osservazioni sui concetti fondamentali della probabilità 615

§5. Processi deterministici e processi casuali 620

§6. Processi casuali di tipo markoviano 625

Letture consigliate 628

Note al capitolo XI 629

L'approssimazione delle funzioni 632

§1. Introduzione 632

§2. I polinomi d'interpolazione 635

§3. Approssimazione degli integrali definiti 642

§4. Il concetto di migliore approssimazione uniforme di Čebyšev 647

§5. I polinomi di Čebyšev che si discostano meno da zero 650

§6. Il teorema di Weierstrass e il legame fra migliore approssimazione e derivabilità 652

§7. Serie di Fourier 655

§8. Approssimazione in media quadratica 661

Letture consigliate 664

Note al capitolo XII 665

Metodi di approssimazione e tecniche di calcolo 667

§1. Approssimazione e metodi numerici 667

§2. I più semplici strumenti ausiliari di calcolo 681

Letture consigliate 689

Note al capitolo XIII 690

Le macchine calcolatrici elettroniche 692

§1. Scopi e principi fondamentali dei calcolatori elettronici 692

§2. Programmazione e codifica per i calcolatori elettronici veloci 696

§3. Principi tecnici delle unità di un calcolatore veloce 706

§4. Prospettive di sviluppo e impiego dei calcolatori elettronici 719

Letture consigliate 725

Note al capitolo XIV 726

Teoria delle funzioni di una variabile reale 728

§1. Introduzione 728

§2. Insiemi 729

§3. Numeri reali[1] 737

§4. Insiemi di punti 743

§5. Misura degli insiemi 749

§6. L'integrale di Lebesgue 754

Letture consigliate 758

Note al capitolo XV 759

Algebra lineare 761

§1. Oggetto e strumenti dell'algebra lineare 761

§2. Spazi lineari 771

§3. Sistemi di equazioni lineari 781

§4. Trasformazioni lineari 792

§5. Forme quadratiche 799

§6. Funzioni di matrici e alcune loro applicazioni 805

Letture consigliate 809

Nota al capitolo XVI 809

Geometria non euclidea 811

§1. Storia del postulato di Euclide 812

§2. La soluzione di Lobačevskij 815

§3. La geometria di Lobačevskij 818

§4. Il significato concreto della geometria di Lobačevskij 826

§5. Gli assiomi della geometria e la loro verifica nel modello 833

§6. Teorie geometriche autonome nate dalla geometria euclidea 839

§7. Spazi multidimensionali 845

§8. Ampliamento del campo della geometria 857

§9. Geometria riemanniana 867

§10. Geometria astratta e spazio reale 877

Letture consigliate 884

Note al capitolo XVII 885

Topologia 890

§1. L'oggetto della topologia 890

§2. Le superfici 894

§3. Le varietà 898

§4. Il metodo combinatorio 901

§5. I campi vettoriali 908

§6. Lo sviluppo della topologia 912

§7. Spazi metrici e topologici 914

Letture consigliate 917

Note al capitolo XVIII 918

Analisi funzionale 920

§1. Lo spazio a n dimensioni 921

§2. Lo spazio di Hilbert: uno spazio a infinite dimensioni 924

§3. Sviluppi in sistemi ortogonali di funzioni 929

§4. Le equazioni integrali 935

§5. Operatori lineari e ulteriori sviluppi dell'analisi funzionale 941

Letture consigliate 948

Note al capitolo XIX 949

Gruppi e altri sistemi algebrici 951

§1. Introduzione 951

§2. Simmetria e trasformazioni 952

§3. Gruppi di trasformazioni 961

§4. Gruppi di Fedorov (gruppi cristallografici) 970

§5. Gruppi di Galois 978

§6. Concetti fondamentali della teoria generale dei gruppi 981

§7. Gruppi continui 987

§8. Gruppi fondamentali 989

§9. Rappresentazioni e caratteri dei gruppi 994

§10. La teoria generale dei gruppi 998

§11. Numeri ipercomplessi 998

§12. Algebre associative 1005

§13. Algebre di Lie 1012

§14. Anelli 1014

§15. Reticoli 1017

§16. Altri sistemi algebrici 1019

Letture consigliate 1020

Note al capitolo XX 1020

Indice analitico 1022

Indice dei nomi 1036

Prefazione all’edizione russa

La matematica, nata nell'antichità dalle esigenze della vita quotidiana, si è sviluppata fino a costituire un vastissimo sistema di discipline assai diverse. Come le altre scienze, essa riflette le leggi del mondo materiale che ci circonda e rappresenta un potente strumento per conoscere e dominare la natura. Tuttavia, l'elevato grado di astrazione che le è proprio rende i suoi rami più recenti relativamente inaccessibili ai non specialisti. Questo carattere astratto ha dato origine, fin dall'antichità, a concezioni idealistiche circa la sua indipendenza dal mondo materiale.

Nel preparare quest'opera, gli autori si sono proposti di far conoscere a una cerchia sufficientemente ampia dell'intelligencija sovietica le diverse discipline matematiche, i loro contenuti e metodi, i fondamenti sui quali poggiano e le vie lungo le quali si sono sviluppate.

Come requisito minimo, si è supposta nel lettore soltanto la conoscenza della matematica della scuola secondaria; i volumi differiscono però tra loro per l'accessibilità degli argomenti trattati. Chi desidera accostarsi per la prima volta agli elementi della matematica superiore potrà leggere con profitto i primi capitoli; per comprendere pienamente le parti successive, sarà invece necessario aver studiato almeno in parte i corrispondenti manuali. Una comprensione approfondita dell'opera nel suo insieme sarà possibile soltanto a chi abbia già una certa familiarità con le applicazioni dell'analisi matematica, vale a dire con il calcolo differenziale e integrale. Per questi lettori, ossia per gli insegnanti di matematica e i docenti delle discipline ingegneristiche e delle scienze naturali, sarà particolarmente importante leggere i capitoli dedicati ai rami più recenti della matematica.

Naturalmente, nei limiti di una sola opera non è stato possibile esaurire tutta la ricchezza neppure dei risultati più fondamentali della ricerca matematica; una certa libertà nella scelta degli argomenti era quindi inevitabile. Nelle sue linee generali, tuttavia, il libro offrirà un'idea dello stato attuale della matematica, delle sue origini e del suo probabile sviluppo futuro. Per questa ragione esso si rivolge, in una certa misura, anche a chi conosce già gran parte degli argomenti esposti. Potrà forse contribuire a superare una certa ristrettezza di vedute che talvolta si riscontra in alcuni dei nostri giovani matematici.

I singoli capitoli sono stati scritti da autori diversi, i cui nomi figurano nell'indice. Nel suo insieme, però, il libro è il risultato di un lavoro collettivo. Il piano generale, la scelta degli argomenti e le successive stesure dei singoli capitoli sono stati sottoposti a discussione comune e migliorati attraverso un vivace scambio di opinioni. Mediante discussioni organizzate, matematici di varie città dell'Unione Sovietica hanno potuto formulare numerose e preziose osservazioni sulla prima versione del testo. Gli autori hanno tenuto conto dei loro pareri e suggerimenti.

Gli autori di alcuni capitoli hanno inoltre partecipato direttamente alla stesura definitiva di altri: la parte introduttiva del capitolo II è stata scritta essenzialmente da B. N. Delone, mentre D. K. Faddeev ha contribuito attivamente alla preparazione dei capitoli IV e XX.

Hanno collaborato anche persone diverse dagli autori dei singoli capitoli: il § 4 del capitolo XIV è stato scritto da L. V. Kantorovič, il § 6 del capitolo VI da O. A. Ladyženskaja e il § 5 del capitolo X da A. G. Postnikov; O. A. Olejnik ha lavorato al testo del capitolo V e Ju. V. Prochorov a quello del capitolo XI.

Alcune sezioni dei capitoli I, II, VII e XVII sono state scritte da V. A. Zalgaller. La revisione del testo definitivo è stata curata da V. A. Zalgaller e V. S. Videnskij, con la collaborazione di T. V. Rogožkinaja e A. P. Leonova.

La maggior parte delle illustrazioni è stata preparata da E. P. Sen'kin.

Mosca, 1956

Il comitato di redazione

CAPITOLO I

Uno sguardo generale sulla matematica

A. D. Aleksandrov

Una presentazione adeguata di una scienza non può limitarsi a informazioni dettagliate, per quanto ampie. Deve anche offrire una visione corretta della sua natura essenziale nel suo insieme. Questo capitolo si propone di dare un quadro generale della natura della matematica. A tale scopo non occorre addentrarsi nei particolari delle teorie matematiche più recenti: la matematica elementare e la storia della disciplina forniscono già una base sufficiente per trarre conclusioni generali.

§ 1. I caratteri distintivi della matematica

1. Astrazioni, dimostrazioni, applicazioni.

Anche con una conoscenza superficiale della matematica, è facile riconoscere alcuni tratti caratteristici: la sua astrattezza, la sua precisione, il suo rigore logico, il carattere indiscutibile delle sue conclusioni e, infine, la gamma eccezionalmente ampia delle sue applicazioni.

L'astrattezza della matematica è facile da vedere. Operiamo con numeri astratti senza preoccuparci di come metterli in relazione in ogni caso con oggetti concreti. A scuola studiamo la tavola pitagorica astratta, cioè una tabella per moltiplicare un numero astratto per un altro, non un numero di ragazzi per un numero di mele, o un numero di mele per il prezzo di una mela.

Allo stesso modo in geometria consideriamo, ad esempio, le linee rette e non i fili tesi, essendo il concetto di linea geometrica ottenuto per astrazione da tutte le altre proprietà, tranne che dall'estensione in una direzione. Più in generale, il concetto di figura geometrica è il risultato dell'astrazione da tutte le proprietà degli oggetti reali tranne la loro forma spaziale e le loro dimensioni.

Astrazioni di questo tipo sono caratteristiche di tutta la matematica. Il concetto di numero intero e quello di figura geometrica sono solo due dei suoi concetti più antichi e più elementari. Sono stati seguiti da una massa di altri, troppo numerosi per essere descritti, che si estendono ad astrazioni come numeri complessi, funzioni, integrali, differenziali, funzionali, spazi n-dimensionali e persino a dimensione infinita, e così via. Queste astrazioni, per così dire ammassate le une sulle altre, sono arrivate a un tale grado di generalizzazione che perdono apparentemente ogni legame con la vita quotidiana e il "comune mortale" non ne capisce nulla oltre al semplice fatto che "tutto questo è incomprensibile".

In realtà, ovviamente, le cose non stanno affatto così. Sebbene il concetto di spazio *n*-dimensionale sia senza dubbio estremamente astratto, ha tuttavia un contenuto completamente reale, che non è molto difficile da comprendere. Nel presente libro sarà nostro compito sottolineare e chiarire il contenuto concreto di concetti astratti come quelli sopra menzionati, affinché il lettore possa convincersi che essi sono tutti connessi con la vita reale, sia nella loro origine che nelle loro applicazioni.

Ma l'astrazione non è proprietà esclusiva della matematica; è caratteristica di ogni scienza, anche di tutta l'attività mentale in generale. Di conseguenza, l'astrattezza dei concetti matematici non fornisce di per sé una descrizione completa del carattere peculiare della matematica.

Le astrazioni della matematica si distinguono per tre caratteristiche. In primo luogo si occupano soprattutto dei rapporti quantitativi e delle forme spaziali, astraendoli da tutte le altre proprietà degli oggetti. In secondo luogo, esse si presentano in una sequenza di gradi crescenti di astrazione, andando molto oltre in questa direzione rispetto alle astrazioni di altre scienze. Illustreremo queste due caratteristiche in dettaglio più avanti, utilizzando come esempi le nozioni fondamentali di numero e figura. Infine, ed è ovvio, la matematica in quanto tale si muove quasi tutta nel campo dei concetti astratti e delle loro interrelazioni. Mentre lo scienziato naturale ricorre costantemente agli esperimenti per provare le sue asserzioni, il matematico impiega solo argomentazioni e calcoli.

È vero che anche i matematici si avvalgono costantemente, per aiutarli nella scoperta dei loro teoremi e dei loro metodi, di modelli e di analoghi fisici, e ricorrono a diversi esempi del tutto concreti. Questi esempi servono come fonte effettiva della teoria e come mezzo per scoprirne i teoremi, ma nessun teorema appartiene definitivamente alla matematica finché non è stato rigorosamente dimostrato mediante un ragionamento logico. Se un geometra, presentando un teorema appena scoperto, lo dimostrasse mediante modelli e si limitasse a tale dimostrazione, nessun matematico ammetterebbe che il teorema è stato dimostrato. L'esigenza di una dimostrazione di un teorema è ben nota nella geometria delle scuole superiori, ma pervade tutta la matematica. Potremmo misurare gli angoli alla base di mille triangoli isosceli con estrema precisione, ma un procedimento del genere non ci fornirebbe mai una dimostrazione matematica del teorema che gli angoli alla base di un triangolo isoscele sono uguali. La matematica esige che questo risultato sia dedotto dai concetti fondamentali della geometria, i quali attualmente, in considerazione del fatto che la geometria è sviluppata su basi rigorose, sono precisamente formulati negli assiomi. E così è in ogni caso. Dimostrare un teorema significa per il matematico dedurlo con un ragionamento logico dalle proprietà fondamentali dei concetti che ricorrono in quel teorema. In questo modo non solo i concetti ma anche i metodi della matematica sono astratti e teorici.

I risultati della matematica si distinguono per un alto grado di rigore logico, e un ragionamento matematico è condotto con tale scrupolosità da renderlo incontestabile e del tutto convincente per chiunque lo capisca. La scrupolosità e la persuasività delle dimostrazioni matematiche sono già ben note in un corso di scuola superiore. Le verità matematiche sono infatti il modello di ciò che è del tutto incontestabile. Non per niente si dice

“chiaro come due più due fa quattro”. Qui viene introdotta la relazione “due più due fa quattro” come immagine stessa dell'inconfutabile e incontestabile.

Ma il rigore della matematica non è assoluto; è in un processo di sviluppo continuo; i principi della matematica non si sono solidificati una volta per tutte, ma vivono di vita propria e possono anche essere oggetto di dispute scientifiche.

In ultima analisi, la vitalità della matematica nasce dal fatto che i suoi concetti e risultati, nonostante tutta la loro astrattezza, hanno origine, come vedremo, nel mondo reale e trovano applicazione molto varia nelle altre scienze, nell'ingegneria e in tutte le attività pratiche della vita quotidiana; rendersi conto di questo è il prerequisito più importante per comprendere la matematica.

L’eccezionale ampiezza delle applicazioni è un altro carattere distintivo della matematica.

In primo luogo, usiamo costantemente, quasi ogni ora, nell'industria e nella vita privata e sociale, i concetti e i risultati più diversi della matematica, senza pensarci affatto; ad esempio, usiamo l'aritmetica per calcolare le nostre spese o la geometria per calcolare la superficie di un appartamento. Naturalmente, le regole qui sono molto semplici, ma dovremmo ricordare che in alcuni periodi dell'antichità rappresentavano le conquiste matematiche più avanzate dell'epoca.

In secondo luogo, la tecnologia moderna sarebbe impossibile senza la matematica. Probabilmente non esiste un solo processo tecnico che possa essere portato a termine senza calcoli più o meno complicati; e la matematica gioca un ruolo molto importante nello sviluppo di nuovi rami della tecnologia.

È vero, infine, che ogni scienza, in misura maggiore o minore, fa un uso essenziale della matematica. Le “scienze esatte”, la meccanica, l'astronomia, la fisica e in larga misura la chimica, esprimono le loro leggi, come ogni scolaretto sa, mediante formule e fanno largo uso di strumenti matematici nello sviluppo delle loro teorie. Il progresso di queste scienze sarebbe stato del tutto impossibile senza la matematica. Per questo motivo le esigenze della meccanica, dell'astronomia e della fisica hanno sempre esercitato un'influenza diretta e decisiva sullo sviluppo della matematica.

Nelle altre scienze la matematica gioca un ruolo minore, ma anche qui trova importanti applicazioni. Naturalmente, nello studio di fenomeni così complessi come quelli che si verificano in biologia e in sociologia, il metodo matematico non può svolgere lo stesso ruolo che, diciamo, in fisica. In tutti i casi, ma soprattutto dove i fenomeni sono più complicati, dobbiamo tenere presente, se non vogliamo perderci in un gioco insensato con le formule, che l'applicazione della matematica è significativa solo se i fenomeni concreti sono già stati fatti oggetto di una teoria profonda. In un modo o nell’altro, la matematica trova applicazione in quasi tutte le scienze, dalla meccanica all’economia politica.

Ricordiamo alcune applicazioni particolarmente brillanti della matematica nelle scienze esatte e nella tecnologia.

Il pianeta Nettuno, uno dei più distanti del Sistema Solare, fu scoperto nel 1846 sulla base di calcoli matematici. Analizzando alcune irregolarità nel movimento di Urano, gli astronomi Adams e Leverrier giunsero alla conclusione che queste irregolarità fossero causate dall'attrazione gravitazionale di un altro pianeta. Leverrier calcolò in base alle leggi della meccanica esattamente dove doveva trovarsi questo pianeta, e un osservatore al quale comunicò i suoi risultati lo avvistò nel suo telescopio nella esatta posizione indicata da Leverrier. Questa scoperta fu un trionfo non solo per la meccanica e l'astronomia, e in particolare per il sistema di Copernico, ma anche per le capacità del calcolo matematico.

Un altro esempio, non meno impressionante, è stata la scoperta delle onde elettromagnetiche. Il fisico inglese Maxwell, generalizzando le leggi dei fenomeni elettromagnetici stabilite sperimentalmente, riuscì ad esprimere queste leggi sotto forma di equazioni. Da queste equazioni dedusse, con metodi puramente matematici, che le onde elettromagnetiche potevano esistere e che dovevano propagarsi con la velocità della luce. Sulla base di questo risultato propose la teoria elettromagnetica della luce, che venne poi sviluppata e approfondita in ogni direzione. Inoltre, i risultati di Maxwell portarono alla ricerca di onde elettromagnetiche di origine puramente elettrica, derivanti ad esempio da una carica oscillante. Queste onde furono effettivamente scoperte da Hertz. Poco dopo, A. S. Popov, scoprendo mezzi per eccitare, trasmettere e ricevere le oscillazioni elettromagnetiche, le rese disponibili per un'ampia gamma di applicazioni e gettò così le basi per l'intera tecnologia della radio. Nella scoperta della radio, ormai possesso comune di tutti, un ruolo importante ebbero i risultati di una deduzione puramente matematica.

Quindi dall'osservazione, come ad esempio della deflessione di un ago magnetico da parte di una corrente elettrica, la scienza procede alla generalizzazione, alla teoria dei fenomeni, alla formulazione di leggi e alla loro espressione matematica. Da queste leggi derivano nuove deduzioni e, infine, la teoria trova realizzazione nella pratica, che a sua volta fornisce nuovi potenti impulsi per lo sviluppo della teoria.

È particolarmente degno di nota il fatto che anche le costruzioni più astratte della matematica, che nascono all'interno della scienza stessa, senza alcuna motivazione immediata da parte delle scienze naturali o della tecnologia, abbiano tuttavia applicazioni fruttuose. Ad esempio, i numeri immaginari vennero alla luce per la prima volta nell'algebra, e per molto tempo il loro significato nel mondo reale rimase incompreso, circostanza indicata dal loro stesso nome. Ma quando intorno al 1800 fu data loro un'interpretazione geometrica (vedi capitolo IV, §2), i numeri immaginari si stabilirono saldamente in matematica, dando origine alla vasta teoria delle funzioni di una variabile complessa, cioè di una variabile della forma $x + y\sqrt{-1}$. Questa teoria delle funzioni "immaginarie" di una variabile "immaginaria" si è rivelata tutt'altro che immaginaria, ma piuttosto un mezzo molto pratico per risolvere problemi tecnologici. Con questa teoria si dimostrano così i risultati fondamentali di N. E. Jukovski sulla portanza dell'ala di un aeroplano. La stessa teoria è utile, ad esempio, nella soluzione dei problemi riguardanti la filtrazione dell'acqua sotto una diga, problemi la cui importanza è evidente nell'attuale periodo di costruzione di grandi centrali idroelettriche.

Un altro esempio, altrettanto impressionante, è fornito dalla geometria non euclidea[1], nata dagli sforzi, estesi per 2000 anni a partire dai tempi di Euclide, per dimostrare l'assioma delle parallele, un problema di interesse puramente matematico. Lo stesso N. I. Lobačevskij, il fondatore della nuova geometria, fu attento a definire la sua geometria "immaginaria", poiché non riusciva a vederne alcun significato nel mondo reale, sebbene fosse fiducioso che un tale significato alla fine sarebbe stato trovato. I risultati della sua geometria apparivano alla maggioranza dei matematici non solo "immaginari" ma addirittura inimmaginabili e assurdi. Tuttavia, le sue idee gettarono le basi per un nuovo sviluppo della geometria, vale a dire la creazione di teorie di vari spazi non euclidei; e queste idee successivamente divennero la base della teoria generale della relatività, in cui l'apparato matematico consiste in una forma di geometria non euclidea dello spazio quadridimensionale. Così le costruzioni astratte della matematica, che per lo meno sembravano incomprensibili, si rivelarono un potente strumento per lo sviluppo di una delle più importanti teorie della fisica. Allo stesso modo, nell'attuale teoria dei fenomeni atomici, nella cosiddetta meccanica quantistica, si fa un uso essenziale di molti concetti e teorie matematiche estremamente astratte, come ad esempio il concetto di spazio a dimensione infinita.

Non è necessario fare ulteriori esempi, poiché abbiamo già mostrato con sufficiente chiarezza che la matematica trova ampia applicazione nella vita quotidiana, nella tecnologia e nella scienza; nelle scienze esatte e nei grandi problemi della tecnologia trovano applicazione anche quelle teorie che nascono all'interno della matematica stessa. Questa è una delle peculiarità caratteristiche della matematica, insieme alla sua astrattezza, al rigore e alla incontrovertibilità dei suoi risultati.

2. La natura essenziale della matematica.

Nel discutere queste particolarità della matematica siamo stati lontani dallo spiegare la sua essenza; piuttosto abbiamo semplicemente sottolineato i suoi caratteri esteriori. Il nostro compito ora è spiegare la natura essenziale di questi tratti caratteristici. A tal fine sarà necessario rispondere, almeno, alle seguenti domande:

Cosa riflettono questi concetti matematici astratti? In altre parole, qual è il vero oggetto della matematica?

Perché i risultati astratti della matematica appaiono così convincenti e i suoi concetti iniziali così ovvi? In altre parole, su quale fondamento poggiano i metodi della matematica?

Perché, nonostante tutta la sua astrattezza, la matematica trova così vasta applicazione e non si rivela un semplice gioco con le astrazioni? In altre parole, come spiegare il significato della matematica?

Infine, quali forze portano all'ulteriore sviluppo della matematica, permettendole di unire l'astrattezza all'ampiezza di applicazione? Qual è la base per la sua continua crescita?

Nel rispondere a queste domande formeremo un quadro generale del contenuto della matematica, dei suoi metodi, del suo significato e del suo sviluppo; cioè ne comprenderemo l'essenza.

Idealisti e metafisici non solo cadono in confusione nel tentativo di rispondere a queste domande fondamentali, ma arrivano al punto di distorcere completamente la matematica, capovolgendola letteralmente. Pertanto, vedendo l'estrema astrattezza e persuasività dei risultati matematici, l'idealista immagina che la matematica provenga dal pensiero puro.

In realtà la matematica non offre il minimo sostegno all'idealismo o alla metafisica. Ce ne convinceremo tentando, a grandi linee, di rispondere alle domande elencate sull'essenza della matematica. Per un chiarimento preliminare di queste questioni è sufficiente esaminare i fondamenti dell'aritmetica e della geometria elementare, ai quali ci rivolgiamo ora.

§ 2. L’aritmetica

1. Il concetto di numero intero.

Il concetto di numero (per ora parliamo solo di numeri interi positivi), anche se oggi ci è così familiare, è stato elaborato molto lentamente. Ciò si può vedere dal modo in cui contavano i vari popoli che fino a tempi recenti sono rimaste ad un livello relativamente primitivo di vita sociale. Presso alcuni di essi non esistevano nomi per numeri superiori a due o tre; tra gli altri, il conteggio andava oltre ma finiva dopo pochi numeri, dopodiché si diceva semplicemente “molti” o “innumerevoli”. Solo gradualmente tra i vari popoli si accumulò un patrimonio di nomi per i numeri chiaramente distinti.

All'inizio questi popoli non avevano idea di cosa fosse un numero, anche se potevano, a modo loro, giudicare la grandezza dell'uno o dell'altro insieme di oggetti che incontravano nella loro vita quotidiana. Dobbiamo concludere che il numero veniva da loro direttamente percepito come una proprietà inseparabile di un insieme di oggetti, proprietà che però non distinguevano chiaramente. Siamo così abituati a contare che difficilmente possiamo immaginare questo stato di cose, ma è possibile capirlo.[2]

Al livello immediatamente superiore il numero appare già come proprietà di una collezione di oggetti, ma non è ancora distinto dalla collezione come “numero astratto”, come numero in generale, non connesso con oggetti concreti. Ciò è evidente dai nomi dei numeri presso alcuni popoli, come “mano” per cinque e “uomo intero” per venti. Qui cinque è da intendersi non in astratto ma semplicemente nel senso di “tante quante sono le dita di una mano”, venti è “tante quante sono le dita delle mani e dei piedi di un uomo” e così via. In modo del tutto analogo, alcuni popoli non avevano il concetto di “nero”, “duro” o “circolare”. Per dire che un oggetto è nero, lo hanno confrontato ad esempio con un corvo, e per dire che c'erano cinque oggetti, hanno confrontato direttamente questi oggetti con una mano. In questo modo avvenne anche che per diversi tipi di oggetti venissero usati nomi diversi per i numeri; alcuni numeri per contare le persone, altri per contare le barche e così via, fino a dieci diversi tipi di numeri. Qui non abbiamo numeri astratti, ma semplicemente una sorta di “denominazione”, riferita solo a un determinato tipo di oggetti. Presso gli altri popoli in genere non esistevano nomi distinti per i numeri, come ad esempio nessuna parola per “tre”, sebbene potessero dire “tre uomini” o “in tre posti”, e così via.

Allo stesso modo tra di noi diciamo abbastanza facilmente che questo o quell’oggetto è nero ma molto più raramente parliamo di “nerezza” in sé, che è un concetto più astratto.[3]

Il numero di oggetti in una data collezione è una proprietà della collezione, ma il numero stesso, in quanto tale, il “numero astratto”, è una proprietà astratta dalla collezione concreta e considerata semplicemente in sé, come “l’nerezza” o la “durezza”. Come l’nerezza è la proprietà comune a tutti gli oggetti del colore del carbone, così il numero “cinque” è la proprietà comune di tutte le collezioni contenenti tanti oggetti quante sono le dita di una mano. In questo caso l'uguaglianza dei due numeri si stabilisce mediante un semplice confronto: prendiamo un oggetto dalla collezione, pieghiamo un dito e contiamo in questo

modo fino alla fine della collezione. Più in generale, accoppiando gli oggetti di due collezioni è possibile, senza fare alcun uso di numeri, stabilire se le collezioni contengono o meno lo stesso numero di oggetti. Ad esempio, se gli ospiti stanno prendendo posto al tavolo, potranno facilmente, senza contare, far capire alla padrona di casa che ha dimenticato un posto, poiché un ospite rimarrà senza posto.

In questo modo è possibile dare la seguente definizione di numero: ogni singolo numero come “due”, “cinque” e così via, è quella proprietà delle collezioni di oggetti che è comune a tutte le collezioni i cui oggetti possono essere messi in corrispondenza biunivoca tra loro e che è diversa per quelle collezioni per le quali tale corrispondenza è impossibile. Per scoprire questa proprietà e distinguerla chiaramente, cioè per formare il concetto di un numero determinato e dargli il nome “sei”, “dieci”, ecc., è stato necessario confrontare molti insiemi di oggetti. Per innumerevoli generazioni gli uomini hanno ripetuto la stessa operazione milioni di volte e hanno così scoperto i numeri e le relazioni tra loro.

2. Relazioni tra i numeri interi.

Le operazioni con i numeri sono nate a loro volta come riflesso delle relazioni tra oggetti concreti. Ciò è osservabile anche nei nomi dei numeri. Ad esempio, presso alcuni indiani d'America il numero "ventisei" viene pronunciato come "sopra le due decine metto un sei", il che riflette chiaramente un metodo concreto di conteggio degli oggetti. L'addizione di numeri corrisponde a mettere insieme o unire due o più raccolte, ed è altrettanto facile vedere il significato concreto di sottrazione, moltiplicazione e divisione. La moltiplicazione in particolare è nata in larga misura, sembra chiaro, dall'abitudine di contare le collezioni uguali: cioè per due, per tre e così via.

Nel contare, gli uomini non solo scoprirono e assimilarono i rapporti tra i singoli numeri, come per esempio che due e tre fanno cinque, ma stabilirono anche gradualmente alcune leggi generali. Con l'esperienza pratica si è scoperto che una somma non dipende dall'ordine degli addendi e che il risultato del conteggio di un dato insieme di oggetti non dipende dall'ordine in cui avviene il conteggio, fatto che si riflette nell'identità essenziale dei numeri “ordinali” e “cardinali”: primo, secondo, terzo e uno, due, tre. In questo modo i numeri non apparivano separati e indipendenti, ma interconnessi tra loro.

Alcuni numeri sono espressi in termini di altri nei loro stessi nomi e nel modo in cui sono scritti. Pertanto, “venti” denota “due (volte) dieci”; in francese ottanta è “quattro-venti” (quatre-vingt), novanta è “quattro-venti-dieci”; e i numeri romani VIII, IX indicano che $8 = 5 + 3$, $9 = 10 - 1$.

In generale, non sono sorti solo numeri separati, ma un sistema di numeri con relazioni e regole reciproche.

L'oggetto dell'aritmetica è proprio questo, il sistema dei numeri con le sue reciproche relazioni e regole.[4] Il singolo numero astratto di per sé non ha proprietà tangibili e in generale c'è ben poco da dire al riguardo. Se ci interroghiamo, ad esempio, sulle proprietà del numero sei, notiamo che $6 = 5 + 1$, $6 = 3 \cdot 2$, che 6 è un divisore di 30 e così via. Ma qui il numero 6 è sempre connesso con altri numeri; infatti, le proprietà di un dato numero consistono proprio

nelle sue relazioni con gli altri numeri.[5] Di conseguenza, è chiaro che ogni operazione aritmetica determina una connessione o relazione tra numeri. Pertanto l'oggetto dell'aritmetica sono le relazioni tra i numeri. Ma queste relazioni sono le immagini astratte delle reali relazioni quantitative tra insiemi di oggetti; si può quindi dire che l'aritmetica è la scienza dei rapporti quantitativi effettivi considerati astrattamente, cioè puramente come rapporti. L'aritmetica, come vediamo, non è nata dal pensiero puro, come sostengono gli idealisti, ma è il riflesso di determinate proprietà delle cose reali; nasce dalla lunga esperienza pratica di molte generazioni.

3. Simboli per i numeri.

Man mano che la vita sociale diventava più estesa e complicata, poneva problemi più ampi. Non solo era necessario prendere nota del numero degli oggetti di un insieme e comunicarlo ad altri, necessità che aveva già portato alla formulazione del concetto di numero e ai nomi dei numeri, ma diventava essenziale imparare a contare collezioni sempre più grandi, di animali in un gregge, di oggetti di scambio, di giorni prima di una data fissa, e così via, e comunicare ad altri i risultati del conteggio. Questa situazione richiedeva assolutamente un miglioramento dei nomi e anche dei simboli dei numeri.

L'introduzione dei simboli per i numeri, avvenuta a quanto pare non appena si cominciò a scrivere, giocò un ruolo importante nello sviluppo dell'aritmetica. Inoltre, fu il primo passo verso i segni e le formule matematiche in generale. Il secondo passo, consistente nell'introduzione dei segni per le operazioni aritmetiche e di una designazione letterale per l'incognita (x), è stato compiuto molto più tardi.

Il concetto di numero, come ogni altro concetto astratto, non ha un'immagine immediata; non può essere esibito ma può solo essere concepito nella mente. Ma il pensiero è formulato nel linguaggio, sicché senza nome non può esistere concetto. Anche il simbolo è un nome, solo che non è orale ma scritto e si presenta alla mente sotto forma di immagine visibile. Ad esempio, se dico “sette”, cosa ti immagini? Probabilmente non un insieme di sette oggetti di un tipo o dell’altro, ma piuttosto il simbolo “7”, che forma una sorta di struttura tangibile per il numero astratto “sette”. Inoltre il numero 18273 è molto più difficile da pronunciare che da scrivere e non può essere rappresentato con precisione sotto forma di un insieme di oggetti. In questo modo avvenne, anche se solo dopo un certo lasso di tempo, che i simboli dessero origine alla concezione di numeri così grandi che non avrebbero mai potuto essere scoperti mediante l'osservazione diretta o l'enumerazione. Con la comparsa dello Stato, fu necessario riscuotere le tasse, radunare ed equipaggiare un esercito, e così via, il che richiese operazioni con numeri molto grandi.

L'importanza dei simboli per i numeri consiste quindi innanzitutto nel fatto che essi forniscono una semplice incarnazione del concetto di numero astratto.[6] Questo è il ruolo delle designazioni matematiche in generale: forniscono un’incarnazione di concetti matematici astratti. Pertanto + denota addizione, x denota un numero incognito, a un numero arbitrario e così via. In secondo luogo i simboli dei numeri forniscono un mezzo particolarmente semplice per effettuare operazioni su di essi. Tutti sanno quanto sia più facile “calcolare su carta” che “a mente”. I segni e le formule matematiche hanno in generale questo

vantaggio: ci permettono di sostituire una parte dei nostri argomenti con calcoli, con qualcosa di quasi meccanico. Del resto, se un calcolo viene trascritto, possiede già una certa evidenza; tutto è visibile, tutto è controllabile e tutto è definito da regole precise. Come esempi si potrebbe citare l'addizione in colonna o qualsiasi trasformazione algebrica come "passare all'altro lato dell'equazione con cambio di segno". Da quanto detto risulta chiaro che senza simboli adeguati per i numeri l'aritmetica non avrebbe potuto fare molti progressi. A maggior ragione è vero che la matematica contemporanea sarebbe impossibile senza i suoi segni e le sue formule particolari.

È ovvio che il metodo estremamente comodo di scrivere i numeri in uso oggi non avrebbe potuto essere elaborato tutto in una volta. Fin dall'antichità sono apparsi presso vari popoli, fin dagli inizi della loro cultura, vari simboli per i numeri, che erano molto diversi da quelli nostri contemporanei non solo nell'aspetto generale ma anche nei principi in base ai quali venivano scelti. Ad esempio, il sistema decimale non era usato ovunque, e presso gli antichi babilonesi esisteva un sistema in parte decimale e in parte sessagesimale. La tabella 1 fornisce alcuni dei simboli dei numeri tra i vari popoli. In particolare vediamo che gli antichi greci, e più tardi anche i russi, facevano uso delle lettere per designare i numeri. I nostri simboli "arabi" contemporanei e, più in generale, il nostro metodo di formazione dei numeri, furono portati dall'India in Europa dagli arabi nel X secolo e vi si radicarono saldamente nel corso dei secoli successivi.

Tavola 1

I simboli numerici presso diversi popoli

	Slavic		Chinese			
	Cyrillic	Glagolitic	Ancient	Commercial	Scientific	Greek
0				〇	〇	
1			一	丨	Ⅰ	
2			二	𠄌	Ⅱ	
3			三	𠄍	Ⅲ	
4			四	〤	Ⅳ	
5			五	𠌶	Ⅴ	
6			六	上	Ⅵ	
7			七	𠂇	Ⅶ	
8			八	𠂆	Ⅷ	
9			九	𠂉	Ⅸ	
10			十	十	Ⅹ	
20			廿	卄	Ⅹⅹ	
30			卅	卅	Ⅹⅹⅹ	
100			百	𠀆	Ⅽ	
1000			千	千	Ⅿ	

Intestazioni: slavo (cirillico, glagolitico); cinese (antico, commerciale, scientifico); greco.

Tavola 1 (continuazione)

I simboli numerici presso diversi popoli

			Egyptian			
	Arabic	Georgian	Hieroglyphic	Hieratic	Roman	Mayan
0						
1	١	ა	𓏺	𓏺	I	•
2	٢	ბ	𓏺𓏺	𓏺𓏺	II	••
3	٣	გ	𓏺𓏺𓏺	𓏺𓏺𓏺	III	•••
4	٤	დ	𓏺𓏺𓏺𓏺	𓏺𓏺𓏺𓏺	IV	••••
5	٥	ე	𓐁	𓐁	V	—
6	٦	ვ	𓐁𓏺	𓐁𓏺	VI	—•
7	٧	ზ	𓐁𓏺𓏺	𓐁𓏺𓏺	VII	—••
8	٨	თ	𓐁𓏺𓏺𓏺	𓐁𓏺𓏺𓏺	VIII	—•••
9	٩	ი	𓐁𓏺𓏺𓏺𓏺	𓐁𓏺𓏺𓏺𓏺	IX	—••••
10	١٠	კ	𓎆	𓎆	X	=
20	٢٠	ლ	𓎆𓎆	𓎆𓎆	XX	
30	٣٠	მ	𓎆𓎆𓎆	𓎆𓎆𓎆	XXX	
100	١٠٠	ნ	𓍢	𓍢	C	
1000	١٠٠٠	ო	𓆼	𓆼	M	

Intestazioni: arabo; georgiano; egizio (geroglifico, ieratico); romano; maya.

La prima particolarità del nostro sistema è che è un sistema decimale. Ma questa non è una questione di grande importanza, poiché sarebbe stato del tutto possibile utilizzare, ad esempio, un sistema duodecimale introducendo simboli speciali per dieci e undici. La peculiarità più importante del nostro sistema di designazione dei numeri è che è "posizionale"; cioè che la stessa cifra ha un significato diverso a seconda della sua posizione. Ad esempio, nel 372 la cifra 3 indica le centinaia e la cifra 7 le decine. Questo metodo di scrittura non solo è conciso e semplice, ma rende i calcoli molto facili. I numeri romani erano sotto questo aspetto molto meno comodi, lo stesso numero 372 essendo scritto nella forma CCCLXXII; è un compito molto laborioso moltiplicare insieme due grandi numeri scritti in numeri romani.

La scrittura posizionale dei numeri esige che in un modo o nell'altro si prenda nota di ogni ordine di unità che è stata omessa, poiché altrimenti confonderemo, ad esempio, trentuno con trecentouno. Nella posizione della posizione omessa dobbiamo mettere uno zero, distinguendo così 301 e 31. In una forma rudimentale, lo zero appare già nelle scritture cuneiformi tardo babilonesi, ma la sua introduzione sistematica fu una conquista degli indiani:[7] Essa ha permesso loro di procedere ad un sistema di scrittura completamente posizionale proprio come lo abbiamo oggi.

Ma in questo modo anche lo zero divenne un numero ed entrò nel sistema dei numeri. Di per sé lo zero non è nulla; nella lingua sanscrita dell'antica India si chiama proprio così: "vuoto" (çūnga); ma in connessione con altri numeri, lo zero acquista contenuto e proprietà ben note; ad esempio, un numero arbitrario più zero è lo stesso numero, oppure quando un numero arbitrario viene moltiplicato per zero diventa zero.

4. La teoria dei numeri come branca della matematica pura.

Torniamo all'aritmetica degli antichi. I testi più antichi conservati da Babilonia e dall'Egitto risalgono al secondo millennio a.C. Questi e successivi testi contengono vari problemi aritmetici con le loro soluzioni, tra cui alcuni che oggi appartengono all'algebra, come la risoluzione di equazioni quadratiche e persino cubiche, oppure lo studio delle progressioni; il tutto presentato, ovviamente, sotto forma di problemi concreti ed esempi numerici. Presso i Babilonesi troviamo anche alcune tavole di quadrati, cubi e reciproci. C'è da supporre che cominciassero già a formarsi interessi matematici che non erano immediatamente collegati a problemi pratici.

In ogni caso l'aritmetica era ben sviluppata nell'antica Babilonia e in Egitto. Non si trattava però ancora di una teoria matematica dei numeri ma piuttosto di una raccolta di soluzioni di vari problemi e di regole di calcolo. Proprio così l'aritmetica viene insegnata ancora oggi nelle nostre scuole elementari e viene compresa da tutti coloro che non sono particolarmente interessati alla matematica. Ciò è perfettamente legittimo, ma l'aritmetica in questa forma non è ancora una teoria matematica. Non esistono teoremi generali sui numeri.

Il passaggio all'aritmetica teorica è avvenuto gradualmente.

Come è stato sottolineato, l'esistenza dei simboli ci permette di operare con numeri così grandi che è impossibile visualizzarli come raccolte di oggetti o arrivarvi contando in

successione a partire dal numero uno. Presso le tribù primitive si contavano singolarmente i numeri fino a 3, 10, 100 e così via, ma dopo questi venivano i “molti” indefiniti. In contrasto con questa situazione, l’uso dei simboli per i numeri permise ai cinesi, ai babilonesi e agli egiziani di arrivare a decine di migliaia e perfino a milioni. Fu in questa fase che si notò la possibilità di estendere indefinitamente la serie dei numeri, anche se non sappiamo quanto presto questa possibilità fu chiaramente percepita. Persino Archimede (287-212 a.C.) nel suo straordinario saggio “Arenario” si dedicò a descrivere un metodo per nominare un numero maggiore del numero di granelli di sabbia sufficienti a riempire la “sfera delle stelle fisse”. Quindi la possibilità di nominare e scrivere un numero del genere richiedeva ancora a suo tempo una spiegazione dettagliata.

Nel III secolo a.C. i Greci avevano chiaramente riconosciuto due idee importanti: primo, che la sequenza dei numeri poteva essere estesa indefinitamente e secondo, che non solo era possibile operare con numeri dati arbitrariamente ma discutere i numeri in generale, formulare e dimostrare teoremi generali su di essi. Questa idea rappresenta la generalizzazione di un'immensa quantità di esperienze precedenti con i numeri concreti, da cui sono emerse le regole e i metodi per il ragionamento generale sui numeri. Si è verificato un passaggio a un livello di astrazione più elevato: dai numeri dati separati (anche se astratti) al numero in generale, a qualsiasi numero possibile.

Dal semplice processo di contare gli oggetti uno per uno si passa al processo illimitato di formazione dei numeri aggiungendo uno al numero già formato. La sequenza dei numeri è considerata continuabile indefinitamente e con essa entra nella matematica la nozione di infinito. Naturalmente non possiamo infatti, mediante il processo di aggiunta di uno, procedere arbitrariamente lungo la sequenza dei numeri: chi potrebbe arrivare fino a un milione di milioni, che è quasi quaranta volte il numero di secondi in mille anni? Ma non è questo il punto; il processo di addizione di unità, il processo di formazione di collezioni di oggetti arbitrariamente grandi è in linea di principio illimitato, per cui esiste la possibilità di continuare la sequenza dei numeri oltre ogni limite. Il fatto che nella pratica il conteggio sia limitato non è rilevante; se ne fa un'astrazione. È di questa successione indefinitamente prolungata che si occupano i teoremi generali sui numeri.

I teoremi generali su qualsiasi proprietà di un numero arbitrario contengono già in forma implicita infinite asserzioni sulle proprietà di numeri separati e sono quindi qualitativamente più ricchi di qualsiasi asserzione particolare che potrebbe essere verificata per numeri specifici. È per questa ragione che i teoremi generali devono essere dimostrati con argomenti generali partendo dalla regola fondamentale per la formazione della sequenza dei numeri. Qui si coglie una peculiarità profonda della matematica: la matematica ha come oggetto non solo i rapporti quantitativi dati, ma tutti i rapporti quantitativi possibili e quindi l'infinito.

Nei famosi “Elementi” di Euclide, scritti nel III secolo a.C., troviamo già teoremi generali sui numeri interi, in particolare il teorema secondo cui esistono numeri primi arbitrariamente grandi.[8]

L’aritmetica si trasforma così nella teoria dei numeri. Si è ormai allontanata dai problemi particolari e concreti per raggiungere la regione dei concetti e degli argomenti astratti. È

diventata parte della matematica “pura”. Più precisamente, questo fu il momento della nascita della matematica pura stessa con i tratti caratteristici discussi nella nostra prima sezione. Bisogna naturalmente prendere atto del fatto che la matematica pura è nata contemporaneamente dall'aritmetica e dalla geometria e che nelle regole generali dell'aritmetica si trovano già alcuni rudimenti dell'algebra, materia che più tardi dall'aritmetica si separerà. Ma di questo parleremo più avanti.

Resta ora da riassumere le nostre conclusioni fin qui, poiché abbiamo ora tracciato, seppure in modo molto affrettato, il processo attraverso il quale l'aritmetica teorica è nata dal concetto di numero.

5. L'essenza dell'aritmetica.

Poiché la nascita dell’aritmetica teorica è parte della nascita della matematica, possiamo ragionevolmente aspettarci che le nostre conclusioni sull’aritmetica gettino luce sulle nostre precedenti domande riguardanti la matematica in generale. Ricordiamo queste domande, in particolare nella loro applicazione all'aritmetica.

1. Come sono nati i concetti astratti dell’aritmetica e che cosa riflettono del mondo reale?

A questa domanda rispondono le osservazioni precedenti sulla nascita dell'aritmetica. I suoi concetti corrispondono alle relazioni quantitative delle collezioni di oggetti. Questi concetti sono nati per astrazione, come risultato dell'analisi e della generalizzazione di un'immensa quantità di esperienza pratica. Sono sorti gradualmente; prima vennero i numeri legati a oggetti concreti, poi i numeri astratti e infine il concetto di numero in generale, di ogni numero possibile. Ciascuno di questi concetti è stato reso possibile dalla combinazione di esperienza pratica e precedenti concetti astratti. Questa, tra l'altro, è una delle leggi fondamentali della formazione dei concetti matematici: essi vengono realizzati da una serie di successive astrazioni e generalizzazioni, ciascuna basata su una combinazione di esperienza con precedenti concetti astratti. La storia dei concetti dell'aritmetica mostra quanto sia sbagliata la concezione idealistica secondo cui essi derivino dal «pensiero puro», dall'«intuizione innata», dalla «contemplazione delle forme a priori», o simili.

2. Perché le conclusioni dell'aritmetica sono così convincenti e inalterabili?

La storia risponde anche a questa domanda per noi. Vediamo che le conclusioni dell'aritmetica sono state elaborate lentamente e gradualmente; riflettono l'esperienza accumulata nel corso di un numero inimmaginabile di generazioni e in questo modo si sono fissate saldamente nella mente dell'uomo. Si sono fissate anche nel linguaggio: nei nomi dei numeri, nei loro simboli, nella ripetizione costante delle stesse operazioni con i numeri, nella loro costante applicazione alla vita quotidiana. È in questo modo che hanno acquisito chiarezza e certezza. Anche i metodi del ragionamento logico hanno la stessa fonte. L'essenziale qui non è solo il fatto che essi possono essere ripetuti a piacimento, ma la loro solidità e chiarezza, che hanno in comune con i rapporti tra le cose del mondo reale, rapporti che si riflettono nei concetti dell'aritmetica e nelle regole della deduzione logica.

Questo è il motivo per cui i risultati dell'aritmetica sono così convincenti; le sue conclusioni derivano logicamente dai suoi concetti fondamentali, ed entrambi, i metodi della logica e i concetti dell'aritmetica, sono stati elaborati e fissati saldamente nella nostra coscienza da tremila anni di esperienza pratica, sulla base di uniformità oggettive nel mondo che ci circonda.

3. Perché l'aritmetica ha un'applicazione così ampia nonostante l'astrattezza dei suoi concetti?

La risposta è semplice. I concetti e le conclusioni dell'aritmetica, che generalizzano un'enorme quantità di esperienza, riflettono in forma astratta quei rapporti che nel mondo reale si incontrano costantemente e ovunque. È possibile contare gli oggetti in una stanza, le stelle, le persone, gli atomi e così via. L'aritmetica considera alcune delle loro proprietà generali, astraendo da tutto ciò che è particolare e concreto, ed è proprio perché considera solo queste proprietà generali che le sue conclusioni sono applicabili a tanti casi. La possibilità di un'ampia applicazione è garantita dall'astrattezza stessa dell'aritmetica, anche se qui è importante che questa astrazione non sia vuota ma derivi da una lunga esperienza pratica. Lo stesso vale per tutta la matematica e per qualsiasi concetto o teoria astratta. Le possibilità di applicazione di una teoria dipendono dall'ampiezza del materiale originale che essa generalizza.

Allo stesso tempo ogni concetto astratto, in particolare il concetto di numero, è limitato nel suo significato a causa della sua stessa astrattezza. In primo luogo, quando applicato a qualsiasi oggetto concreto, riflette solo un aspetto dell'oggetto e quindi ne dà solo un'immagine incompleta. Quante volte accade, ad esempio, che i meri dati numerici dicano ben poco sull'essenza della questione. In secondo luogo, i concetti astratti non possono essere applicati ovunque senza determinate condizioni limite; è impossibile applicare l'aritmetica a problemi concreti senza prima convincerci che la sua applicazione abbia un senso nel caso particolare. Se parliamo ad esempio di addizione e ci limitiamo a unire gli oggetti nel pensiero, allora naturalmente non è avvenuto alcun cambiamento negli oggetti stessi. Ma se all'unione effettiva degli oggetti applichiamo l'addizione, se effettivamente mettiamo insieme gli oggetti, ad esempio gettandoli in una pila o disponendoli su un tavolo, in questo caso non avviene solo un'addizione astratta, ma anche un processo reale. Questo processo non consiste semplicemente nell'addizione aritmetica, e in generale può essere addirittura impossibile eseguirlo. Ad esempio, l'oggetto gettato in una pila potrebbe rompersi; gli animali selvatici, se messi insieme, possono dilaniarsi a vicenda; i materiali messi insieme possono entrare in una reazione chimica: un litro di acqua e un litro di alcol versati insieme hanno prodotto non 2, ma 1,9 litri di miscela per effetto della parziale soluzione dei liquidi; e così via.

Se sono necessari altri esempi, sono facili da produrre.

Per dirla in breve, la verità è concreta; ed è particolarmente importante ricordare questo fatto rispetto alla matematica, proprio per la sua astrattezza.

4. Infine, l'ultima questione che abbiamo sollevato riguardava le forze che hanno portato allo sviluppo della matematica.

Anche per l'aritmetica la risposta a questa domanda è chiara dalla sua storia. Abbiamo visto come le persone nel mondo reale imparassero a contare e a elaborare il concetto di numero, e come la vita pratica, ponendo problemi più difficili, necessitasse di simboli per i numeri. In una parola, le forze che portarono allo sviluppo dell'aritmetica furono le esigenze pratiche della vita sociale. Questi bisogni pratici e il pensiero astratto che ne deriva esercitano l'uno sull'altro una costante interazione. I concetti astratti forniscono di per sé uno strumento prezioso per la vita pratica e vengono costantemente migliorati dalla loro stessa applicazione. L'astrazione da tutto ciò che non è essenziale svela il nocciolo della questione e garantisce il successo nei casi in cui un ruolo decisivo è giocato dalle proprietà e dalle relazioni individuate e preservate dall'astrazione; vale a dire, nel caso dell'aritmetica, dalle relazioni quantitative.

Inoltre, la riflessione astratta spesso va oltre le esigenze immediate di un problema pratico. Così il concetto di numeri così grandi come un milione o un miliardo è nato sulla base di calcoli pratici, ma è sorto prima della necessità pratica di farne uso. Ci sono molti esempi simili nella storia della scienza; è sufficiente ricordare i numeri immaginari menzionati prima. Questo è solo un caso particolare di un fenomeno noto a tutti, ovvero l'interazione tra esperienza e pensiero astratto, tra pratica e teoria.

§ 3. La geometria

1. Il concetto di figura geometrica.

La storia dell'origine della geometria è essenzialmente simile a quella dell'aritmetica. Anche i primi concetti e informazioni geometrici risalgono alla preistoria e risultano anch'essi dall'attività pratica.

I primi uomini presero le forme geometriche dalla natura. Il cerchio e la falce della luna, la superficie liscia di un lago, la rettilineità di un raggio di luce o di un albero ben proporzionato esistevano molto prima dell'uomo stesso e si presentavano costantemente alla sua osservazione. Ma nella natura stessa i nostri occhi raramente incontrano linee realmente rette, triangoli o quadrati precisi, ed è chiaro che la ragione principale per cui gli uomini gradualmente elaborarono una concezione di queste figure è che la loro osservazione della natura era attiva, nel senso che, per soddisfare i loro bisogni pratici, fabbricavano oggetti di forma sempre più regolare. Costruirono abitazioni, tagliarono pietre, recintarono appezzamenti di terreno, tesero le corde degli archi, modellarono le loro ceramiche di argilla, le portarono alla perfezione e di conseguenza formarono l'idea che una pentola è curva, ma una corda tesa è diritta. Insomma, prima hanno dato forma alla loro materia e solo poi hanno riconosciuto la forma come ciò che è impresso sulla materia e può quindi essere considerata in sé, come un'astrazione dalla materia. Riconoscendo la forma dei corpi, l'uomo ha potuto perfezionare la sua opera e così elaborare in modo ancora più preciso la nozione astratta di forma. Così l'attività pratica serviva come base per i concetti astratti della geometria. È stato necessario fabbricare migliaia di oggetti con bordi diritti, tendere migliaia di fili, tracciare sul suolo un gran numero di linee rette, prima che gli uomini potessero formarsi una nozione chiara della linea retta in generale, come quella qualità che è comune a tutti questi casi particolari. Oggigiorno impariamo presto a tracciare una linea retta, poiché siamo circondati da oggetti con bordi diritti che sono il risultato di una lavorazione, ed è solo per questo che già nell'infanzia ci formiamo un'idea chiara della linea retta. Allo stesso modo è nata dall'attività pratica la nozione di grandezza geometrica, di lunghezza, di area e di volume. Le persone misuravano lunghezze, determinavano distanze, stimavano a occhio l'area delle superfici e i volumi dei corpi, tutto per i loro scopi pratici. Fu così che furono scoperte le leggi generali più semplici, le prime relazioni geometriche: per esempio, che l'area di un rettangolo è uguale al prodotto delle lunghezze dei suoi lati. È utile che l'agricoltore sia a conoscenza di tale relazione, per poter stimare la superficie che ha seminato e di conseguenza il raccolto che si può aspettare.

Vediamo così che la geometria è nata dall'attività pratica e dai problemi della vita quotidiana. A questo proposito l'antico studioso greco Eudemo di Rodi scrisse quanto segue: "La geometria fu scoperta dagli Egiziani in seguito alla misurazione della terra. Questa misurazione era per loro necessaria a causa delle inondazioni del Nilo, che costantemente spazzavano via i loro confini.[9] Non c'è nulla di straordinario nel fatto che questa scienza, come le altre, sia nata dai bisogni pratici degli uomini. Ogni conoscenza che nasce da circostanze imperfette tende a perfezionarsi. Essa nasce dalle impressioni dei sensi ma

diventa gradualmente oggetto della nostra contemplazione e infine entra nel regno dell'intelletto".

Naturalmente la misurazione della terra non fu l'unico problema che condusse gli antichi verso la geometria. Dai testi frammentari sopravvissuti è possibile farsi un'idea dei vari problemi degli antichi egizi e babilonesi e dei loro metodi per risolverli. Uno dei testi egiziani più antichi risale al 1700 a.C. Si tratta di un manuale di istruzioni per "segretari" (funzionari reali), scritto da un certo Ahmes. Contiene una raccolta di problemi sul calcolo della capacità di contenitori e magazzini, dell'area delle appezzamenti di terreno, delle dimensioni dei lavori di sterro, e così via.

Gli egiziani e i babilonesi sapevano determinare le aree e i volumi più semplici, conoscevano con notevole esattezza il rapporto tra la circonferenza e il diametro di un cerchio, e forse erano anche in grado di calcolare la superficie di una sfera; in una parola, possedevano già un notevole bagaglio di conoscenze geometriche. Ma per quanto ne sappiamo, non erano ancora in possesso della geometria come scienza teorica con teoremi e dimostrazioni. Come l'aritmetica dell'epoca, la geometria era fondamentalmente un insieme di regole dedotte dall'esperienza. Inoltre, la geometria in generale non era distinta dall'aritmetica. I problemi geometrici erano allo stesso tempo problemi di calcolo aritmetico.

Nel VII secolo a.C. la geometria passò dall'Egitto alla Grecia, dove fu ulteriormente sviluppata dai grandi filosofi materialisti, Talete, Democrito e altri. Un notevole contributo alla geometria fu dato anche dai successori di Pitagora, i fondatori di una scuola filosofico-religiosa idealistica.

Lo sviluppo della geometria ha preso la direzione di raccogliere nuovi fatti e di chiarire i loro rapporti reciproci. Queste relazioni furono gradualmente trasformate in deduzioni logiche di certe proposizioni di geometria da certe altre. Ciò ebbe due risultati: primo, il concetto di teorema geometrico e la sua dimostrazione; e in secondo luogo, il chiarimento di quelle proposizioni fondamentali da cui si possono dedurre le altre, cioè gli assiomi.

In questo modo la geometria si sviluppò gradualmente in una teoria matematica.

È noto che esposizioni sistematiche di geometria apparvero in Grecia già nel V secolo a.C., ma non si sono conservate, per l'ovvia ragione che furono tutte soppiantate dagli "Elementi" di Euclide (III secolo a.C.). In quest'opera la geometria veniva presentata come un sistema così ben formato che nulla di essenziale fu aggiunto alle sue fondamenta fino al tempo di N. I. Lobačevskij, più di duemila anni dopo. Il noto testo scolastico di Kiselev, come i libri scolastici di tutto il mondo, non rappresentava nelle sue edizioni più antiche altro che una rielaborazione popolare di Euclide. Pochissimi altri libri al mondo hanno avuto una vita così lunga come gli "Elementi" di Euclide, questa perfetta creazione del genio greco. Naturalmente la matematica ha continuato a progredire e la nostra comprensione dei fondamenti della geometria si è notevolmente approfondita; tuttavia gli "Elementi" di Euclide divennero, e restano in gran parte, il modello di un libro di matematica pura. Riunendo le conquiste dei suoi predecessori, Euclide presentò la matematica del suo tempo come una scienza teorica indipendente; cioè la presentò essenzialmente come la intendiamo oggi.

2. La natura essenziale della geometria.

La storia della geometria porta alle stesse conclusioni di quella dell'aritmetica. Vediamo che la geometria è nata dalla vita pratica e che la sua trasformazione in teoria matematica ha richiesto un tempo immenso.

La geometria opera con “corpi geometrici” e figure; studia i loro reciproci rapporti dal punto di vista della grandezza e della posizione. Ma un corpo geometrico non è altro che un corpo reale considerato unicamente dal punto di vista della sua forma spaziale,[10] astraendo da tutte le altre sue proprietà come densità, colore o peso. La figura geometrica è un concetto ancora più generale, poiché in questo caso è possibile astrarre anche dall'estensione spaziale; quindi una superficie ha solo due dimensioni, una linea, solo una dimensione, e un punto, nessuna. Un punto è il concetto astratto dell’estremo di una linea, di una posizione definita al limite della precisione in modo tale da non avere più parti. È così che tutti questi concetti vengono definiti da Euclide.

Così la geometria ha per oggetto le forme e le relazioni spaziali dei corpi reali, separate dalle loro altre proprietà e considerate dal punto di vista puramente astratto. È proprio questo alto livello di astrazione che distingue la geometria dalle altre scienze che indagano anche le forme spaziali e le relazioni dei corpi. In astronomia, ad esempio, si studiano le reciproche posizioni dei corpi, ma sono i corpi celesti reali; nella geodesia si studia la forma della terra, nella cristallografia la forma dei cristalli, e così via. In tutte queste altre scienze si studia la forma e la posizione dei corpi concreti nella loro dipendenza da altre proprietà dei corpi.

Questa astrazione conduce necessariamente al metodo puramente teorico della geometria; non è più possibile compiere esperimenti con linee rette prive di spessore, con “forme pure”. L’unica possibilità è utilizzare un ragionamento logico, traendo alcune conclusioni da altre. Un teorema geometrico deve essere dimostrato col ragionamento, altrimenti non appartiene alla geometria; non si occupa di “forme pure”.

L'evidenza dei concetti fondamentali della geometria, i metodi del ragionamento e la certezza delle loro conclusioni, hanno tutti la stessa fonte dell'aritmetica. Le proprietà dei concetti geometrici, come i concetti stessi, sono state astratte dal mondo che ci circonda. È stato necessario che gli uomini tracciassero innumerevoli linee rette prima di poter prendere come assioma che per ogni due punti è possibile tracciare una linea retta; dovettero muovere vari corpi e applicarli l'uno all'altro in innumerevoli occasioni prima di poter generalizzare la loro esperienza alla nozione di sovrapposizione di figure geometriche e servirsi di questa nozione per la dimostrazione di teoremi, come si fa nei noti teoremi sulla congruenza dei triangoli.

Infine, dobbiamo sottolineare la generalità della geometria. Il volume di una sfera è pari a $4/3\pi\, R^3$ indipendentemente dal fatto che si parli di un vaso sferico, di una sfera d'acciaio, di una stella o di una goccia d'acqua. La geometria può astrarre ciò che è comune a tutti i corpi, perché ogni corpo reale ha forma, dimensioni e posizione più o meno definite rispetto agli altri corpi. Non c’è quindi da meravigliarsi che la geometria trovi applicazione quasi altrettanto estesa dell’aritmetica. Gli operai che misurano le dimensioni di un edificio o leggono un progetto, un artigliere che determina la distanza dal suo bersaglio, un contadino

che misura l'area del suo campo, un ingegnere che stima il volume dei lavori di sterro, tutte queste persone utilizzano gli elementi della geometria. Il navigatore, l'astronomo, il geodeta, l'ingegnere, il fisico, tutti hanno bisogno delle precise conclusioni della geometria.

Un chiaro esempio della soluzione astratto-geometrica di un importante problema della fisica è fornito dalle ricerche del noto cristallografo e geometra E. S. Fedorov. Il problema che si pose di trovare tutte le possibili forme di simmetria per i cristalli è uno dei più fondamentali della cristallografia teorica. Per risolvere questo problema, Fedorov fece un'astrazione da tutte le proprietà fisiche del cristallo, considerandolo solo come un sistema regolare di corpi geometrici "al posto di un sistema di atomi concreti". Il problema divenne quindi quello di trovare tutte le forme di simmetria che potevano esistere in un sistema di corpi geometrici. Questo problema puramente geometrico fu completamente risolto da Fedorov, che trovò tutte le possibili forme di simmetria, 230 in numero. La sua soluzione si è rivelata un importante contributo alla geometria ed è stata la fonte di molte indagini geometriche.

In questo esempio, come in tutta la storia della geometria, rileviamo la forza motrice principale nello sviluppo della geometria. È l'influenza reciproca della vita pratica e del pensiero astratto. Il problema della scoperta delle possibili simmetrie ebbe origine nell'osservazione fisica dei cristalli ma si trasformò in un problema astratto e diede così origine ad una nuova teoria matematica, la teoria dei sistemi regolari, o dei cosiddetti gruppi di Fedorov.[11] Successivamente questa teoria non solo trovò brillante conferma nell'osservazione pratica dei cristalli ma servì anche da guida generale nello sviluppo della cristallografia, dando luogo a nuove indagini, sia nella fisica sperimentale che nella matematica pura.

§ 4. Aritmetica e geometria

1. L'origine delle frazioni nell'interrelazione tra aritmetica e geometria.

Finora abbiamo considerato aritmetica e geometria separatamente l'una dall'altra. La loro mutua relazione, e di conseguenza la più generale interrelazione di tutte le teorie matematiche, è finora sfuggita alla nostra attenzione. Tuttavia questa relazione ha un significato eccezionalmente grande. L'interazione delle teorie matematiche porta a progressi nella matematica stessa e scopre anche un ricco tesoro di relazioni reciproche nel mondo reale riflesso da queste teorie.

Aritmetica e geometria non solo vengono applicate l'una all'altra, ma servono anche come fonti per ulteriori idee, metodi e teorie generali. In ultima analisi, aritmetica e geometria sono le due radici da cui è cresciuta tutta la matematica. La loro reciproca influenza risale al tempo in cui entrambe erano appena nate. Anche la semplice misurazione di una linea rappresenta un'unione di geometria e aritmetica. Per misurare la lunghezza di un oggetto applichiamo ad esso una certa unità di lunghezza e calcoliamo quante volte è possibile farlo; la prima operazione (applicazione) è geometrica, la seconda (conteggio) è aritmetica. Chiunque conta i propri passi lungo una strada già unisce queste due operazioni.

In generale, la misurazione di qualsiasi grandezza combina il conteggio con qualche operazione specifica che è caratteristica di questo tipo di grandezza. Basti pensare alla misurazione di un liquido in un contenitore graduato o alla misurazione di un segmento di tempo contando il numero di oscillazioni di un pendolo.

Ma nel processo di misurazione risulta, in generale, che l'unità scelta non è contenuta nella grandezza misurata un numero intero di volte, per cui un semplice calcolo del numero di unità non è sufficiente. Diventa necessario dividere l'unità di misura per esprimere più accuratamente la grandezza in parti dell'unità; cioè non più per numeri interi ma per frazioni. Fu così che nacquero effettivamente le frazioni, come dimostra l'analisi dei dati storici e di altro tipo. Sono nate dalla divisione e dal confronto di grandezze continue; in altre parole, dalla misurazione. Le prime grandezze da misurare furono geometriche, cioè lunghezze, aree di campi e volumi di liquidi o materiali granulari, così che nella prima comparsa delle frazioni vediamo l'azione reciproca dell'aritmetica e della geometria. Questa interazione porta alla comparsa di un nuovo importante concetto, vale a dire quello di frazioni, come estensione del concetto di numero dai numeri interi ai numeri frazionari (o come dicono i matematici, ai numeri razionali, che esprimono il rapporto tra numeri interi). Le frazioni non sono nate e non potrebbero derivare dalla divisione di numeri interi, poiché solo gli oggetti interi vengono contati con numeri interi. Tre uomini, tre frecce e così via, tutto questo ha senso, ma due terzi di un uomo e anche due terzi di una freccia sono concetti senza senso; anche tre terzi separati di una freccia non uccideranno un cervo, per questo è necessario avere una freccia intera.

2. Grandezze incommensurabili.

Nello sviluppo del concetto di numero, derivante dall'azione reciproca di aritmetica e geometria, la comparsa delle frazioni fu solo il primo passo. La successiva fu la scoperta degli segmenti incommensurabili. Ricordiamo che gli segmenti si dicono incommensurabili se non esiste alcun segmento che possa essere applicato a ciascuno di essi un numero intero di volte o, in altre parole, se il loro rapporto non può essere espresso da una frazione ordinaria; cioè, da un rapporto di numeri interi.

All'inizio semplicemente non si pensava alla questione se ogni segmento potesse essere espresso da una frazione. Se dividendo o misurando un segmento si imbattevano in parti molto piccole, le scartavano semplicemente; in pratica non aveva senso parlare di infinita precisione di misura. Democrito avanzò addirittura l'idea che le figure geometriche siano costituite da atomi di un tipo particolare. Questa nozione, che a nostro avviso sembra piuttosto strana, si è rivelata molto fruttuosa nella determinazione di aree e volumi. Un'area è stata calcolata come la somma di righe costituite da atomi e un volume come la somma degli strati atomici. Fu così, ad esempio, che Democrito trovò il volume di un cono. Un lettore che capisca il calcolo integrale noterà che questo metodo costituisce già il prototipo della determinazione di aree e volumi con i metodi del calcolo integrale. Inoltre, tornando con il pensiero ai tempi di Democrito, bisogna cercare di liberarsi dalle consuete nozioni odierne, che lo sviluppo della matematica ha saldamente fissato nella nostra mente. Al tempo di Democrito le figure geometriche non erano ancora separate da quelle reali nella misura in cui lo sono oggi. Poiché Democrito considerava i corpi reali come costituiti da atomi, naturalmente considerava nella stessa luce anche le figure geometriche.

Ma l'idea che gli segmenti siano costituiti da atomi entra in contraddizione con il teorema di Pitagora, poiché da questo teorema consegue che esistono segmenti incommensurabili. Ad esempio, la diagonale di un quadrato è incommensurabile con il suo lato; in altre parole, il rapporto tra i due non può essere espresso come rapporto tra numeri interi.

Dimostreremo che il lato e la diagonale di un quadrato sono infatti incommensurabili. Se a è il lato e b è la diagonale di un quadrato, allora secondo il teorema di Pitagora $b^2 = a^2 + a^2 = 2a^2$ e quindi

$$\left(\frac{b}{a}\right)^2 = 2.$$

Ma non esiste una frazione tale che il suo quadrato sia uguale a 2. Infatti, se supponiamo che esista, siano p e q numeri interi per i quali

$$\left(\frac{p}{q}\right)^2 = 2,$$

dove possiamo supporre che p e q non abbiano un fattore comune maggiore di uno, altrimenti potremmo semplificare la frazione. Ma se $(p/q)^2 = 2$, allora $p^2 = 2q^2$, e quindi p^2 è divisibile per 2. In questo caso p^2 è anche divisibile per 4, poiché è il quadrato di un numero pari. Quindi $p^2 = 4q_1$; ovvero, $2q^2 = 4q_1$ e $q^2 = 2q_1$. Da ciò consegue che anche q deve essere

divisibile per 2. Ma questo contraddice la supposizione che p e q non abbiano fattori comuni maggiori di uno. Questa contraddizione dimostra che il rapporto b/a non può essere espresso da un numero razionale. La diagonale e il lato di un quadrato sono incommensurabili.

Questa scoperta fece una grande impressione sugli scienziati greci. Al giorno d'oggi, quando siamo abituati ai numeri irrazionali e calcoliamo liberamente con le radici quadrate, l'esistenza di segmenti incommensurabili non ci disturba. Ma nel V secolo a.C. la scoperta di tali segmenti ebbe per i Greci un aspetto completamente diverso. Poiché non conoscevano il concetto di numero irrazionale e non avevano mai scritto un simbolo come $\sqrt{2}$, il risultato precedente indicava che il rapporto tra la diagonale e il lato del quadrato non era rappresentato da alcun numero.

Nell'esistenza di segmenti incommensurabili i Greci scoprirono un profondo paradosso insito nel concetto di continuità, una delle espressioni della contraddizione dialettica compresa nella continuità e nel movimento. Molti importanti filosofi greci consideravano questa contraddizione; tra questi particolarmente noto, per i suoi paradossi, è Zenone l'Eleatico.

I Greci fondarono una teoria dei rapporti degli segmenti, o delle grandezze in generale, che tiene conto dell'esistenza di segmenti incommensurabili;[12] è esposta negli “Elementi” di Euclide, e in forma semplificata viene oggi spiegata nei corsi liceali di geometria. Ma per riconoscere che il rapporto tra un segmento e l'altro (se il secondo segmento è preso come unità di lunghezza, questo rapporto è semplicemente la lunghezza del primo segmento) può essere considerato come un numero, per cui si generalizza il concetto stesso di numero, a questa idea i Greci non poterono elevarsi: il concetto di numero irrazionale semplicemente non ebbe origine tra loro.[13] Questo passo fu fatto più tardi dai matematici orientali; e in generale, una definizione matematicamente rigorosa di numero reale, non dipendente immediatamente dalla geometria, è stata data solo di recente: negli anni settanta del secolo scorso.[14] Il trascorrere di un periodo di tempo così vasto dopo la fondazione della teoria dei rapporti mostra quanto sia difficile scoprire concetti astratti e dar loro una formulazione esatta.

3. Il numero reale.

Nel descrivere il concetto di numero reale, Newton nella sua “Aritmetica universale” scrisse: “per numero non intendiamo tanto un insieme di unità quanto un rapporto astratto tra una certa quantità e un’altra quantità presa come unità”. Questo numero (rapporto) può essere intero, razionale o, se la grandezza data è incommensurabile con l'unità, irrazionale.

Un numero reale nel suo senso originario non è quindi altro che il rapporto di una grandezza rispetto ad un'altra presa come unità; in casi particolari si tratta di un rapporto di segmenti, ma può anche essere un rapporto di aree, pesi e così via.

Di conseguenza, un numero reale è un rapporto di grandezze in generale, considerate in astrazione dalla loro natura concreta.

Come i numeri interi astratti presentano interesse matematico solo nei loro rapporti reciproci, così i numeri reali astratti hanno contenuto e diventano oggetto dell'attenzione matematica solo nei rapporti reciproci nel sistema dei numeri reali.

Nella teoria dei numeri reali, proprio come nell'aritmetica, è prima necessario definire le operazioni sui numeri: addizione, sottrazione, moltiplicazione, divisione, e anche le relazioni espresse da parole come "maggiore di" o "minore di". Queste operazioni e relazioni riflettono le effettive connessioni tra le varie grandezze; ad esempio, l'addizione riflette la messa insieme degli segmenti. Gli inizi delle operazioni con i numeri reali astratti furono compiuti nel Medioevo dai matematici orientali. Successivamente venne gradualmente scoperta la proprietà più importante del sistema dei numeri reali, la sua continuità. Il sistema dei numeri reali è l'immagine astratta di tutti i possibili valori di una grandezza continuamente variabile.

In questo modo, come nel caso analogo dei numeri interi, l'aritmetica dei numeri reali si occupa delle reali relazioni quantitative delle grandezze continue, che studia nella loro forma generale, in completa astrazione da ogni proprietà concreta. Proprio perché i numeri reali hanno a che fare con ciò che è comune a tutte le grandezze continue, hanno un'applicazione così ampia: i valori di varie grandezze, una lunghezza, un peso, l'intensità di una corrente elettrica, l'energia e così via, sono espressi da numeri, e l'interdipendenza o le relazioni tra queste entità si rispecchiano come relazioni tra i loro valori numerici.

Per mostrare come il concetto generale di numeri reali possa servire come base di una teoria matematica, dobbiamo darne la definizione matematica in modo formale. Ciò può essere fatto con vari metodi, ma forse il più naturale è partire dallo stesso processo di misurazione delle grandezze che effettivamente ha portato nella vita pratica a questa generalizzazione del concetto di numero. Parleremo della lunghezza degli segmenti, ma il lettore si accorgerà subito che si potrebbe ragionare esattamente nello stesso modo per qualunque altra grandezza che consenta una suddivisione indefinita.

Supponiamo di voler misurare l'segmento *AB* mediante l'segmento *CD* preso come unità (figura 1).

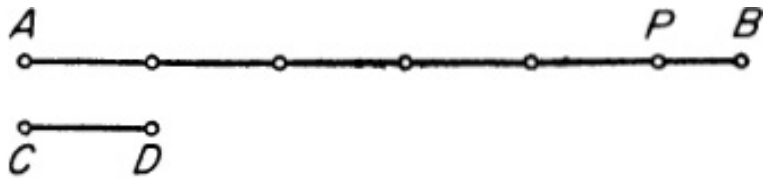


Figura 1.

Riportiamo il segmento *CD* su *AB*, iniziando, per esempio, dal punto *A*, tante volte quante *CD* è contenuto in *AB*. Supponiamo che ciò avvenga n_0 volte. Se del segmento *AB* rimane ancora una parte *PB*, dividiamo *CD* in dieci parti e misuriamo il resto con questi decimi. Supponiamo che nel resto siano contenuti n_1 decimi. Se rimane ancora una parte, dividiamo di nuovo l'unità di misura in dieci parti: dividiamo cioè *CD* in cento parti, ripetiamo la stessa operazione e così via. Il procedimento di misurazione o termina, oppure continua. In entrambi i casi otteniamo che il segmento intero *CD* è contenuto in *AB* n_0 volte; nel resto sono contenuti n_1 decimi, nel resto successivo n_2 centesimi, e così via. In breve, determiniamo il

rapporto tra *AB* e *CD* con precisione crescente: fino ai decimi, ai centesimi e così via. Il rapporto stesso è dunque rappresentato da uno sviluppo decimale con n_0 unità, n_1 decimi e così via:

$$\frac{AB}{CD} = n_0 \cdot n_1 n_2 n_3 \cdots .$$

Questa rappresentazione decimale può essere infinita, corrispondente alla possibilità di aumento indefinito della precisione della misurazione.

Quindi il rapporto tra due segmenti, o tra due grandezze in generale, è sempre rappresentabile da una rappresentazione decimale, finita o infinita. Ma nella rappresentazione decimale non c'è più traccia della grandezza concreta stessa; rappresenta esattamente il rapporto astratto, il numero reale. Quindi un numero reale può essere definito formalmente, se lo si desidera, come rappresentazione decimale finita o infinita.[15]

La nostra definizione sarà completa se diremo cosa intendiamo per operazioni di addizione e così via per le frazioni decimali. Ciò viene fatto in modo tale che le operazioni definite sulle frazioni decimali corrispondano alle operazioni sulle grandezze stesse. Pertanto, quando gli segmenti vengono messi insieme, le loro lunghezze vengono sommate; cioè la lunghezza dell'segmento *AB* + *BC* è uguale alla somma delle lunghezze di *AB* e *BC*. Nel definire le operazioni sui numeri reali, c'è la difficoltà che questi numeri sono rappresentati in generale da frazioni decimali infinite, mentre le regole ben note per queste operazioni si riferiscono a frazioni decimali finite. Una definizione rigorosa delle operazioni per gli sviluppi decimali infiniti può essere fatta nel modo seguente. Supponiamo, ad esempio, di dover sommare i due numeri *a* e *b*. Prendiamo le frazioni decimali corrispondenti fino a una data cifra decimale, diciamo la sesta cifra decimale, e le sommiamo. Otteniamo così la somma *a* + *b* con corrispondente precisione, fino a due milionesimi, poiché gli errori in *a* e *b* possono essere sommati. Siamo quindi in grado di definire la somma di due numeri con un grado di accuratezza arbitrario, e in questo senso la loro somma è completamente definita, sebbene in ogni fase del calcolo sia nota solo con una certa accuratezza. Ma ciò corrisponde all'essenza del caso, poiché anche ciascuna delle grandezze *a* e *b* viene misurata solo con una certa precisione, e il valore esatto di ciascuna delle corrispondenti frazioni infinite si ottiene come risultato di un aumento di precisione prolungato all'infinito. Le relazioni “maggiore di” e “minore di” possono allora essere definite mediante addizione: *a* è maggiore di *b* se esiste una grandezza *c* tale che *a* = *b* + *c*, dove parliamo ovviamente di numeri positivi.

La continuità del sistema dei numeri reali si esprime nel fatto seguente: se i numeri a_1, a_2, ... crescono e i numeri b_1, b_2, ... decrescono, rimanendo però sempre maggiori degli a_l, fra le due successioni esiste sempre un numero *c*. Possiamo rappresentare questo fatto su una retta, mettendone i punti in corrispondenza con i numeri secondo la regola consueta (figura 2).

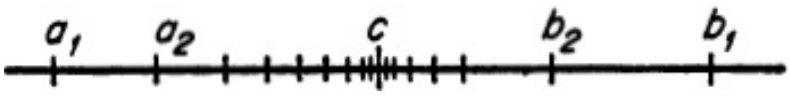


Figura 2.

Qui si vede chiaramente che la presenza del numero *c* e del punto ad esso corrispondente significa l'assenza di un'interruzione nel sistema dei numeri, che è ciò che si intende per loro continuità.

4. Il conflitto degli opposti: concreto e astratto.

Già nell'esempio dell'interazione tra aritmetica e geometria possiamo vedere che lo sviluppo della matematica è un processo di conflitto tra molti elementi contrastanti: il concreto e l'astratto, il particolare e il generale, il formale e il materiale, il finito e l'infinito, il discreto e il continuo, e così via. Proviamo, ad esempio, a tracciare il contrasto tra concreto e astratto nella formazione del concetto di numero reale. Come abbiamo visto, il numero reale riflette un processo di misurazione infinitamente migliorabile o, in termini leggermente diversi, una determinazione assolutamente accurata di una grandezza. Ciò corrisponde al fatto che in geometria consideriamo forme e dimensioni dei corpi idealmente precise, astraendo del tutto dalla mobilità degli oggetti concreti e da una certa indefinitezza nelle loro forme e dimensioni reali; ad esempio, l'segmento misurato (figura 1) era del tutto ideale.

Ma forme geometriche idealmente precise e valori assolutamente precisi per le grandezze rappresentano astrazioni. Nessun oggetto concreto ha una forma assolutamente precisa, né alcuna grandezza concreta può essere misurata con assoluta precisione, poiché non ha nemmeno un valore assolutamente accurato. La lunghezza di un segmento di linea, ad esempio, non ha senso se si cerca di renderla precisa oltre i limiti delle dimensioni atomiche. In ogni caso, quando si superano i limiti ben noti dell'accuratezza quantitativa, appare un cambiamento qualitativo nella grandezza, e in generale essa perde il suo significato originario. Ad esempio, la pressione di un gas non può essere determinata con precisione oltre i limiti dell'impatto di una singola molecola; la carica elettrica cessa di essere continua quando si cerca di renderla precisa oltre la carica di un elettrone e così via. Considerando l'assenza in natura di oggetti dalla forma idealmente precisa, l'affermazione che il rapporto tra la diagonale di un quadrato e il lato è uguale a $\sqrt{2}$ non solo non può essere dedotta con assoluta precisione dalla misurazione immediata, ma non ha nemmeno alcun significato assolutamente preciso per un quadrato concreto.

La conclusione che la diagonale e il lato del quadrato sono incommensurabili deriva, come abbiamo visto, dal teorema di Pitagora. Questa è una conclusione teorica basata sullo sviluppo dei dati dell'esperienza; è il risultato dell'applicazione della logica alle premesse originali della geometria, che sono tratte dall'esperienza.

In questo modo il concetto di segmenti incommensurabili, e tanto più di numeri reali, non è un semplice riflesso immediato dei fatti dell'esperienza ma li oltrepassa. Questo è abbastanza comprensibile. Il numero reale non riflette una grandezza concreta data, ma la grandezza in generale, astraendo da ogni concretezza; in altre parole, riflette ciò che è comune a particolari grandezze concrete. Ciò che è comune a tutti consiste soprattutto nel fatto che il valore della grandezza può essere determinato in modo sempre più preciso; e se astraiamo dalle grandezze concrete, allora il limite di questo possibile aumento di precisione, che dipende dalla natura concreta della grandezza, diventa indefinito e scompare.

In tal modo una teoria matematica delle grandezze, poiché considera le grandezze astraendo dalla loro natura individuale, deve inevitabilmente considerare la possibilità di un'esattezza illimitata per il valore della grandezza e deve con ciò condurre al concetto di numero reale. Allo stesso tempo, poiché riflette solo ciò che è comune alle varie grandezze, la matematica non tiene conto delle peculiarità di ciascuna grandezza individuale.

Poiché la matematica prende in considerazione soltanto proprietà generali, essa opera con le sue astrazioni chiaramente definite in modo del tutto indipendente dai limiti effettivi della loro applicabilità, come deve accadere proprio perché questi limiti sono diversi nei diversi casi particolari. Tali limiti dipendono dalle proprietà concrete dei fenomeni considerati e dai cambiamenti qualitativi che in essi avvengono. Quindi nel fare un'applicazione della matematica è necessario verificare l'effettiva applicabilità della teoria in questione. Considerare la materia come continua e descrivere le sue proprietà mediante grandezze continue è lecito solo se possiamo astrarre dalla sua struttura atomica, e ciò è possibile solo in condizioni ben note.

Tuttavia, i numeri reali rappresentano uno strumento affidabile e potente per l'indagine matematica di grandezze e processi continui reali. La loro teoria si basa sulla pratica, su un immenso campo di applicazioni nella fisica, nella tecnologia e nella chimica. Di conseguenza, la pratica dimostra che il concetto di numero reale riflette correttamente le proprietà generali delle grandezze. Ma questa correttezza non è illimitata; non è possibile considerare la teoria dei numeri reali come qualcosa di assoluto, che consenta uno sviluppo astratto illimitato e in completa separazione dalla realtà. Il concetto stesso di numero reale continua a svilupparsi ed è infatti ancora lungi dall'essere completo.

5. Il conflitto degli opposti: discreto e continuo.

Il ruolo di un altro dei contrasti citati, quello tra il discreto e il continuo, può essere illustrato anche dallo sviluppo del concetto di numero. Abbiamo già visto che le frazioni nascono dalla divisione di grandezze continue.

Su questo tema della divisione c'è una domanda umoristica e straordinariamente istruttiva. La nonna ha comprato tre patate e deve dividerle equamente tra due nipoti. Come può farlo? La risposta è: prepara il purè di patate.

Lo scherzo rivela l'essenza stessa della questione. Gli oggetti separati sono indivisibili nel senso che, una volta diviso, l'oggetto quasi sempre cessa di essere quello che era prima, come risulta chiaramente dall'esempio dei “terzi di uomo” o dei “terzi di freccia”. D'altra parte, grandezze o oggetti continui e omogenei possono essere facilmente divisi e ricomposti senza perdere il loro carattere essenziale. Il purè di patate offre un ottimo esempio di oggetto omogeneo, che di per sé non è diviso in parti ma può tuttavia essere suddiviso in pratica in piccole parti a piacere. Lunghezze, aree e volumi hanno la stessa proprietà. Sebbene siano continui nella loro stessa essenza e non siano effettivamente divisi in parti, tuttavia offrono la possibilità di essere divisi senza limiti.

Qui incontriamo due tipi contrastanti di oggetti: da un lato, gli oggetti indivisibili, separati, discreti; e dall'altro gli oggetti che sono completamente divisibili e tuttavia non sono divisi in

parti ma sono continui. Naturalmente, queste caratteristiche contrastanti sono sempre unite, poiché non esistono oggetti assolutamente indivisibili e oggetti completamente continui. Eppure questi aspetti degli oggetti hanno un'esistenza reale, e spesso accade che un aspetto sia decisivo in un caso e l'altro in un altro.

Astraendo le forme dal loro contenuto, la matematica con questo stesso atto divide nettamente queste forme in due classi, quella discreta e quella continua.

Il modello matematico di un oggetto separato è l'unità, e il modello matematico di un insieme di oggetti discreti è una somma di unità, che è, per così dire, l'immagine della pura discretezza, purificata da ogni altra qualità. D'altronde il modello matematico fondamentale e originale della continuità è la figura geometrica; nel caso più semplice, la linea retta.

Abbiamo quindi davanti a noi due contrasti, la discretezza e la continuità, e le loro immagini matematiche astratte: il numero intero e l'estensione geometrica. La misurazione consiste nell'unione di questi contrasti: Il continuo è misurato da unità separate. Ma le unità inseparabili non bastano; dobbiamo introdurre parti frazionarie dell'unità originale. In questo modo nascono i numeri frazionari e si sviluppa il concetto di numero proprio come risultato dell'unione dei suddetti contrasti.

Poi, a un livello più astratto, è apparso il concetto di segmenti incommensurabili e, di conseguenza, il numero reale come immagine astratta dell'aumento illimitato della precisione nella determinazione di una grandezza. Questo concetto non si è formato immediatamente, e il lungo percorso del suo sviluppo ha portato attraverso molti conflitti tra questi stessi due elementi contrastanti, il discreto e il continuo.

In primo luogo Democrito rappresentò le figure come costituite da atomi e in questo modo ridusse il continuo al discreto. Ma la scoperta di segmenti incommensurabili portò all'abbandono di tale rappresentazione. Dopo questa scoperta le grandezze continue non furono più pensate come costituite da elementi separati, atomi o punti, e non furono rappresentate da numeri, poiché a quel tempo non erano conosciuti numeri diversi dagli interi e dalle frazioni.

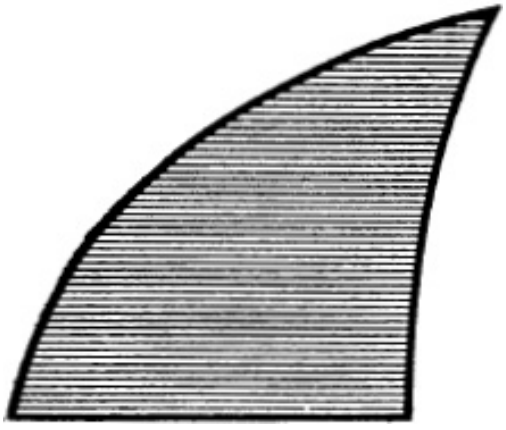


Figura 3.

Il contrasto tra continuo e discreto riapparve in matematica con rinnovata forza nel XVII secolo, quando furono gettate le basi del calcolo differenziale e integrale. Qui era l'infinitesimo ad essere in discussione. In alcune concezioni l'infinitesimo era pensato come una particella reale, “effettivamente” infinitesimale, “indivisibile” di grandezza continua, come gli atomi di Democrito, tranne per il fatto che ora il numero di queste particelle era considerato infinitamente grande. Il calcolo delle aree e dei volumi, o in altre parole l'integrazione, era pensato come la somma di un numero infinito di queste particelle infinitamente piccole. Per area, ad esempio, si intendeva “la somma delle linee da cui è formata” (figura 3). Di conseguenza, il continuo è stato nuovamente ridotto al discreto, ma ora in modo più complicato, a un livello più alto. Ma anche questo punto di vista si rivelò insoddisfacente e, come contrappeso, apparve, sulla base del lavoro di Newton, la nozione di variabili continue, dell’infinitesimo come variabile continua decrescente senza limite. Questa concezione prevalse definitivamente all'inizio del XIX secolo, quando fu fondata la rigorosa teoria dei limiti. Il segmento era ormai concepito non più come un insieme di punti o “indivisibili”, ma come un'estensione, come un mezzo continuo, dove era possibile fissare solo punti separati, valori separati di una grandezza continua. I matematici allora parlarono di “estensione”. Nell'unione del discreto e del continuo è ancora il continuo a dominare.

Ma lo sviluppo dell'analisi esigeva una maggiore precisione nella teoria delle grandezze variabili e soprattutto nella definizione generale di numero reale come valore arbitrario possibile di una grandezza variabile. Negli anni settanta del secolo scorso è nata una teoria dei numeri reali che rappresenta un segmento come un insieme di punti, e corrispondentemente l'intervallo di variazione di una variabile come un insieme di numeri reali. Il continuo era ancora costituito da punti discreti separati e le proprietà della continuità erano ancora espresse nella struttura dell'insieme di punti che lo formavano. Questa concezione portò a immensi progressi in matematica e divenne dominante. Ma ancora una volta si scoprirono profonde difficoltà, che portarono a tentativi di ritornare su un nuovo livello al concetto di pura continuità. Sono stati fatti altri tentativi per modificare il concetto di segmento come insieme di punti. Apparvero nuovi punti di vista per i concetti di numero, variabile e funzione. Lo sviluppo della teoria continua e dobbiamo attendere i suoi ulteriori progressi.

6. Ulteriori risultati dell'interazione tra aritmetica e geometria.

L'interazione tra geometria e aritmetica ha avuto un ruolo anche al di fuori della formazione del concetto di numero reale. La stessa interazione della geometria con l'aritmetica, o più precisamente con l'algebra, si è manifestata anche nella formazione dei numeri negativi e complessi, cioè dei numeri della forma $a + b\sqrt{-1}$. I numeri negativi sono rappresentati dai punti della linea retta a sinistra del punto che rappresenta lo zero. Fu proprio questa rappresentazione geometrica a dare ai numeri immaginari un posto fisso nella matematica; fino a quel momento non erano stati compresi. Apparvero nuovi concetti di grandezza: ad esempio i vettori, che sono rappresentati da segmenti orientati; e tensori, che sono grandezze ancora più generali; anche in questi l'algebra è unita alla geometria.

L'unione di varie teorie matematiche ha sempre giocato un ruolo grande e talvolta decisivo nello sviluppo della matematica. Lo vedremo più avanti nello sviluppo della geometria

analitica, del calcolo differenziale e integrale, della teoria delle funzioni di variabile complessa, della recente cosiddetta analisi funzionale e di altre teorie. Anche nella stessa teoria dei numeri, cioè nello studio dei numeri interi, si applicano con grande successo metodi che dipendono dalla continuità (cioè dall'analisi infinitesimale) e dalla geometria. Questi metodi hanno dato origine ad ampi rami della teoria dei numeri: la “teoria analitica dei numeri” e la “geometria dei numeri”.

Da un determinato punto di vista è possibile considerare i fondamenti della matematica come l'unione di concetti derivanti dalla geometria e dall'aritmetica; cioè dei concetti generali di continuità e di operazioni algebriche (come generalizzazioni di operazioni aritmetiche). Ma non potremo parlare qui di queste difficili teorie. Lo scopo del presente capitolo è stato quello di dare un'idea dell'interazione generale dei concetti, dell'unione e del conflitto tra idee contrastanti in matematica, come illustrato dall'interazione dell'aritmetica e della geometria nello sviluppo del concetto di numero.

§ 5. L’epoca della matematica elementare

1. I quattro periodi della matematica.

Lo sviluppo della matematica non si riduce al semplice accumulo di nuovi teoremi ma comporta cambiamenti qualitativi essenziali. Questi cambiamenti qualitativi non avvengono però in un processo di distruzione o di abolizione delle teorie già esistenti, ma nel loro approfondimento e generalizzazione, in modo da formare teorie più generali, per le quali gli sviluppi precedenti hanno preparato la strada.

Dal punto di vista più generale, possiamo distinguere nella storia della matematica quattro periodi fondamentali, qualitativamente distinti. Naturalmente non è possibile tracciare linee di confine esatte tra queste fasi, poiché i tratti essenziali di ogni periodo sono comparsi più o meno gradualmente, ma le distinzioni tra le fasi ed i passaggi dall'una all'altra sono del tutto chiare.

Il primo stadio (o periodo) è il periodo della nascita della matematica come scienza indipendente e puramente teorica. Essa inizia nei tempi più antichi e si estende fino al V secolo a.C., o forse prima, quando i Greci gettarono le basi della matematica “pura” con la sua connessione logica tra teoremi e dimostrazioni (in quel secolo apparvero, in particolare, esposizioni sistematiche di geometria come gli “Elementi” di Ippocrate di Chio). Questa prima fase fu il periodo della formazione dell'aritmetica e della geometria, nella forma considerata in precedenza. A quel tempo la matematica consisteva in un insieme di regole separate dedotte dall'esperienza e immediatamente connesse con la vita pratica. Queste regole non formavano ancora un sistema logicamente unificato, poiché il carattere teorico della matematica con la sua dimostrazione logica dei teoremi si formò molto lentamente, man mano che si accumulava materiale. Aritmetica e geometria non erano separate ma erano strettamente intrecciate l'una con l'altra.

Il secondo periodo può essere caratterizzato come il periodo della matematica elementare, della matematica delle grandezze costanti; i suoi semplici risultati fondamentali costituiscono ora il contenuto di un corso di scuola superiore. Questo periodo durò quasi 2000 anni e terminò nel XVII secolo con la nascita della matematica “superiore”. È di questo periodo che ci occuperemo più in dettaglio nella presente sezione. Le sezioni successive saranno dedicate al terzo e quarto periodo, cioè alla fondazione e allo sviluppo dell'analisi e al periodo della matematica contemporanea.

2. La matematica in Grecia.

Il periodo della matematica elementare può a sua volta essere diviso in due parti, distinte per il loro contenuto di base: il periodo dello sviluppo della geometria (fino al II secolo d.C.) e il periodo del predominio dell'algebra (dal II al XVII secolo). Rispetto alle condizioni storiche è diviso in tre parti, che possono essere chiamate “greca”, “orientale” e “Rinascimento europeo”. Il periodo greco coincide nel tempo con la fioritura generale della cultura greca, a partire dal VII secolo a.C., per raggiungere il culmine nel III secolo a.C. all'epoca dei grandi geometri dell'antichità, Euclide, Archimede e Apollonio, fino al VI secolo d.C. La matematica, e soprattutto la geometria, conobbero uno sviluppo meraviglioso in Grecia. Conosciamo i nomi e i risultati di numerosi matematici greci, anche se ci sono pervenute solo poche opere autentiche. È da notare che Roma non diede nulla alla matematica, sebbene Roma avesse raggiunto il proprio apice nel I secolo d.C. in un'epoca in cui la scienza della Grecia, conquistata da Roma, era ancora fiorente.

I Greci non solo svilupparono e sistematizzarono la geometria elementare nella misura in cui essa è data negli “Elementi” di Euclide e viene ora insegnata nelle nostre scuole secondarie, ma ottennero risultati notevolmente più elevati. Hanno studiato le sezioni coniche: ellisse, iperbole, parabola; dimostrarono alcuni teoremi relativi agli elementi di quella che viene chiamata geometria proiettiva; guidati dalle esigenze dell'astronomia, elaborarono la geometria sferica (nel I secolo d.C.) e anche gli elementi di trigonometria, e calcolarono le prime tavole dei seni (Ipparco, II secolo a.C. e Claudio Tolomeo, II secolo d.C.);[16] determinarono aree e volumi di alcune figure complicate; ad esempio Archimede trovò l'area del segmento di una parabola dimostrando che è 2/3 dell'area del rettangolo che lo contiene (figura 4). I Greci conoscevano anche il teorema secondo cui tra tutti i corpi con una data superficie la sfera ha il volume maggiore, ma la loro dimostrazione non è stata conservata e probabilmente non era completa. Tale dimostrazione è piuttosto difficile e fu scoperta per la prima volta nel XIX secolo, mediante il calcolo integrale.

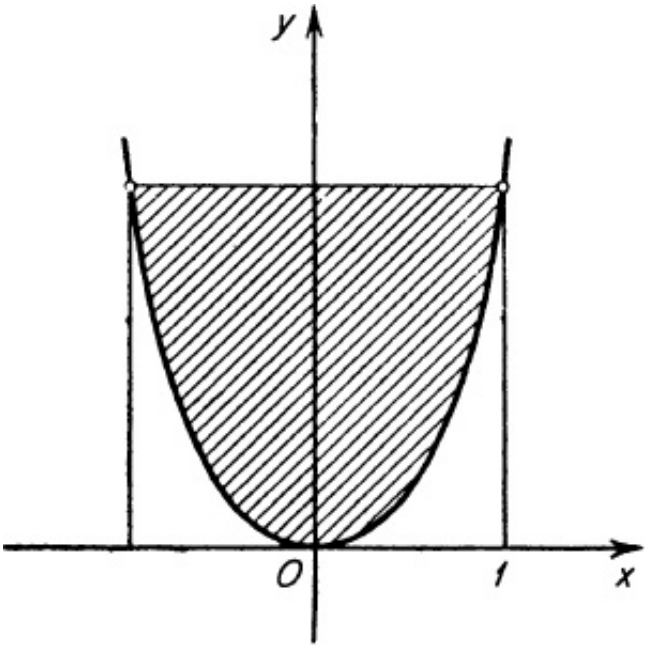


Figura 4.

Anche nell'aritmetica e negli elementi dell'algebra i Greci fecero notevoli progressi. Come accennato in precedenza, gettarono le basi per la teoria dei numeri. Qui rientrano, ad esempio, le loro indagini sui numeri primi (il teorema di Euclide sull'esistenza di un numero infinito di numeri primi e il "crivello" di Eratostene per trovare i numeri primi) e la soluzione di equazioni in numeri interi (Diofanto circa 246-330 d.C.).

Abbiamo già detto che i Greci scoprirono le grandezze irrazionali ma le considerarono geometricamente, come segmenti. Quindi i problemi che oggi affrontiamo algebricamente furono trattati geometricamente dai Greci. Fu in questo modo che risolsero equazioni quadratiche e trasformarono espressioni irrazionali. Ad esempio, l'equazione che oggi scriviamo nella forma $x^2 + ax = b^2$, dicevano così: Trova un segmento x tale che se al quadrato costruito su di esso aggiungiamo un rettangolo costruito sullo stesso segmento e sul segmento dato a, otteniamo un rettangolo di area uguale a un quadrato dato. Questo dominio della geometria durò molto tempo dopo i Greci. Conoscevano anche i metodi (geometrici) per l'estrazione delle radici quadrate e cubiche e le proprietà delle progressioni aritmetiche e geometriche.

In questo modo i Greci erano già in possesso di gran parte del materiale dell'algebra elementare contemporanea ma non, però, dei seguenti elementi essenziali: numeri negativi e zero, numeri irrazionali interamente astratti dalla geometria, e infine un sistema ben sviluppato di simboli letterali. È vero che Diofanto si serviva di simboli letterali per indicare l'incognita e le sue potenze e anche di simboli speciali per addizione, sottrazione e uguaglianza, ma le sue equazioni algebriche erano ancora scritte con coefficienti numerici concreti.

In geometria i Greci raggiunsero quella che oggi chiamiamo matematica "superiore". Archimede si avvalse del calcolo integrale per il calcolo di aree e volumi e Apollonio utilizzò la geometria analitica nelle sue indagini sulle sezioni coniche. Apollonio in realtà fornisce le equazioni di queste curve[17] ma le esprime in linguaggio geometrico. In queste equazioni non appare ancora la nozione generale di costante arbitraria o di grandezza variabile; e i mezzi necessari per esprimere tali concetti, cioè i simboli letterali dell'algebra, compaiono solo in età successiva; soltanto loro potrebbero trasformare tali ricerche in una fonte di nuove teorie, che diventerebbero veramente parte della matematica superiore. I fondatori di queste nuove teorie furono guidati, mille anni dopo, dall'eredità degli scienziati greci; infatti, la "Geometria" di Cartesio (1637), che gettò le basi della geometria analitica, inizia con una selezione di problemi lasciati dai Greci.

Questa è la regola generale. Le vecchie teorie, dando origine a problemi nuovi e profondi, superano, per così dire, se stesse, e richiedono per ulteriori progressi nuove forme e nuove idee. Ma queste forme e idee possono richiedere nuove condizioni storiche per la loro nascita. Nella società antica non esistevano e non potevano esistere le condizioni necessarie per il passaggio alla matematica superiore; esse si realizzarono con lo sviluppo delle scienze naturali in epoca moderna, sviluppo che a sua volta fu condizionato nei secoli XVI e XVII dalle nuove esigenze della tecnica e della manifattura e si legò in tal modo con la nascita e lo sviluppo del capitalismo.

I Greci hanno praticamente esaurito le possibilità della matematica elementare, il che spiega il fatto che il brillante progresso della geometria si è esaurito all'inizio della nostra era e è stato sostituito dalla trigonometria e dall'algebra nelle opere di Tolomeo, Diofanto e altri. In effetti, si può considerare l'opera di Diofanto come l'inizio del periodo in cui l'algebra svolse un ruolo di primo piano. Ma la società degli antichi, già prossima al declino, non era più in grado di far avanzare la scienza in questa nuova direzione.

Va notato che, qualche secolo prima, l'aritmetica aveva già raggiunto un livello elevato in Cina. Gli scienziati cinesi del II e I secolo a.C. descrissero le regole per la soluzione aritmetica di un sistema di tre equazioni di primo grado. È qui che per la prima volta nella storia vengono utilizzati coefficienti negativi e vengono formulate le regole per operare con quantità negative. Ma le soluzioni stesse venivano cercate solo sotto forma di numeri positivi, proprio come più tardi nelle opere di Diofanto. Questi libri cinesi includono anche un metodo per l'estrazione delle radici quadrate e cubiche.

3. Il Medio Oriente.

Con la fine della scienza greca iniziò in Europa un periodo di stagnazione scientifica, il centro dello sviluppo matematico si spostò in India, Asia centrale e paesi arabi.[18] Per un periodo di circa mille anni, dal V al XV secolo, la matematica si sviluppò principalmente in relazione alle esigenze del calcolo, soprattutto in astronomia, poiché i matematici orientali erano per la maggior parte anche astronomi. È vero che non aggiunsero nulla di importante alla geometria greca; in questo campo si conservarono solo per tempi successivi i risultati dei Greci. Ma i matematici indiani, arabi e dell'Asia centrale ottennero immensi successi nei campi dell'aritmetica e dell'algebra.[19]

Come è stato menzionato nel §2, gli indiani hanno inventato il nostro attuale sistema di numerazione. Introdussero anche i numeri negativi, paragonando il contrasto tra numeri positivi e negativi con il contrasto tra proprietà e debito o tra le due direzioni su una linea retta. Infine cominciarono a operare con le grandezze irrazionali esattamente come con quelle razionali, senza rappresentarle geometricamente, a differenza dei Greci. Avevano anche simboli speciali per le operazioni algebriche, inclusa l'estrazione delle radici. Proprio perché gli studiosi indiani e dell'Asia centrale non erano più ostacolati dalla differenza tra le grandezze irrazionali e razionali, furono in grado di superare il "dominio" della geometria, che era caratteristico della matematica greca, e di aprire strade per lo sviluppo dell'algebra contemporanea, libera dalla pesante struttura geometrica in cui era stata costretta dai Greci.

Il grande poeta e matematico Omar Khayyam (circa 1048-1122), e anche il matematico azerbaigiano Nasireddin Tusi (1201-1274), mostrarono chiaramente che ogni rapporto di grandezze, commensurabile o incommensurabile, può essere chiamato numero; nelle loro opere troviamo la stessa definizione generale di numero, sia razionale che irrazionale, come è stata introdotta sopra nella formulazione di Newton, nel §4. La portata di questi risultati diventa particolarmente chiara se ricordiamo che il completo riconoscimento dei numeri negativi e irrazionali fu raggiunto dai matematici europei solo molto lentamente, anche dopo l'inizio del Rinascimento della matematica in Europa. Ad esempio, il celebre matematico francese Viète (1540-1603), al quale l'algebra deve molto, evitava i numeri negativi, e in

Inghilterra le proteste contro di essi durarono fino al XVIII secolo. Questi numeri erano considerati assurdi, poiché erano inferiori a zero, cioè “meno di niente”. Al giorno d'oggi sono diventati familiari, anche se solo sotto forma di temperatura negativa; tutti leggono i giornali e capiscono cosa vuol dire “la temperatura a Mosca è -8°”.

La parola “algebra” deriva dal nome di un trattato del matematico e astronomo Mahommed ibn Musa al-Kharizmi (Mahommed, figlio di Musa, originario di Kharizm), vissuto nel IX secolo. Il suo trattato di algebra si chiamava Al-jebr w’al-muqabala, che significa “trasposizione e rimozione”. Per trasposizione (al-jebr) si intende il trasferimento dei termini negativi all'altro lato di un'equazione, e per rimozione (almuqabala), la cancellazione dei termini uguali su entrambi i lati.

La parola araba “al-jebr” divenne nella trascrizione latina “algebra” e la parola al-muqabala fu scartata, da cui deriva il termine moderno “algebra”.[20]

L'origine di questo termine corrisponde molto bene al contenuto stesso della scienza. L'algebra è fondamentalmente la dottrina delle operazioni aritmetiche considerate formalmente da un punto di vista generale, con l'astrazione dai numeri concreti. I suoi problemi mettono in primo piano le regole formali per la trasformazione delle espressioni e la soluzione delle equazioni. Al-Kharizmi pose sul frontespizio del suo libro i nomi reali di due regole formali più generali, esprimendo in questo modo il vero spirito dell'algebra.

Successivamente, Omar Khayyam definì l'algebra come la scienza della risoluzione delle equazioni. Questa definizione mantenne il suo significato fino alla fine del XIX secolo, quando l'algebra, insieme alla teoria delle equazioni, prese nuove direzioni, cambiando sostanzialmente il suo carattere ma non mutando il suo spirito di generalità come scienza delle operazioni formali.

I matematici dell'Asia centrale trovarono metodi per il calcolo, sia esatto che approssimativo, delle radici di un certo numero di equazioni; scoprirono la formula generale del “binomio di Newton”, sebbene la esprimessero a parole; svilupparono e sistematizzarono notevolmente la scienza della trigonometria e calcolarono tabelle dei seni molto accurate. Queste tabelle furono calcolate, per scopi astronomici, dal matematico Gyaseddin (1427 circa) che lavorava con il famoso astronomo uzbeko Ulug Begh; Gyaseddin inventò anche le frazioni decimali 150 anni prima che venissero reinventate in Europa.

Riassumendo, nel corso del Medioevo in India e nell'Asia centrale l'attuale sistema di numerazione decimale (comprese le frazioni) fu quasi completamente costruito, così come lo furono anche l'algebra elementare e la trigonometria. Nello stesso periodo le conquiste della scienza cinese cominciarono a diffondersi nei paesi vicini; intorno al VI secolo a.C. i cinesi disponevano già di metodi per la soluzione delle più semplici equazioni indeterminate, per calcoli approssimati in geometria e per i primi passi nella soluzione approssimata di equazioni di terzo grado. Essenzialmente le uniche parti del nostro attuale corso di algebra liceale che non erano conosciute prima del XVI secolo erano i logaritmi e i numeri immaginari. Tuttavia non esisteva ancora un sistema di simboli letterali: il contenuto dell’algebra aveva superato la sua forma. Ma la forma era indispensabile: l'astrazione dai numeri concreti e la formulazione

di regole generali richiedevano un corrispondente metodo di espressione; era essenziale avere un modo per denotare numeri arbitrari e le operazioni su di essi. Il simbolismo algebrico è la forma necessaria corrispondente al contenuto dell'algebra. Come nella remota antichità era stato necessario, per poter operare con i numeri interi, inventare dei simboli, così ora, per operare con numeri arbitrari e dare regole generali per il loro utilizzo, è stato necessario elaborare i simboli corrispondenti. Questo compito, iniziato al tempo dei Greci, non fu portato a termine fino al XVII secolo, quando l'attuale sistema di simboli fu finalmente istituito nelle opere di Cartesio e altri.

4. L'Europa rinascimentale.

Al tempo del Rinascimento gli europei conobbero la matematica greca attraverso le traduzioni arabe. I libri di Euclide, Tolomeo e Al-Kharizmi furono tradotti nel XII secolo dall'arabo al latino, la lingua scientifica comune dell'Europa occidentale, e allo stesso tempo il precedente sistema di calcolo, derivato dai Greci e dai Romani, fu gradualmente sostituito dall'attuale metodo indiano, preso in prestito dagli europei dagli arabi.

Fu solo nel XVI secolo che la scienza europea superò finalmente i risultati dei suoi predecessori. Così gli italiani, Tartaglia e Ferrari, risolsero l'equazione generale di terzo grado e, più tardi, l'equazione generale del quarto grado (vedi capitolo IV). Notiamo che, sebbene questi risultati non vengano insegnati a scuola, appartengono, rispetto ai metodi adoperati in essi, all'algebra elementare. All'algebra superiore bisogna però riferire la teoria generale delle equazioni.

Nello stesso periodo cominciarono ad essere utilizzati per la prima volta i numeri immaginari; dapprima ciò avvenne in modo puramente formale, senza una giustificazione logica, che venne fornita soltanto molto più tardi, all'inizio del XIX secolo. Furono elaborati anche i nostri simboli algebrici attuali; in particolare, i simboli letterali furono usati da Viète nel 1591 non solo per le quantità incognite ma anche per quelle date.

Molti matematici hanno preso parte a questo sviluppo dell'algebra. Nello stesso tempo apparvero in Europa le frazioni decimali; furono inventate dallo studioso olandese Stevin, che ne scrisse nel 1585.

Infine, Napier in Gran Bretagna inventò i logaritmi come ausilio nei calcoli astronomici e ne scrisse nel 1614; Briggs calcolò le prime tavole decimali dei logaritmi, che furono pubblicate nel 1624.[21]

Contemporaneamente apparvero in Europa la “teoria delle combinazioni” e la formula generale del “binomio di Newton”;[22] le progressioni erano già note, e così venne completata la struttura dell'algebra elementare. Con ciò finì, all'inizio del XVII secolo, tutto il periodo della matematica delle grandezze costanti, della matematica elementare così come viene insegnata oggi, con poche aggiunte, nelle nostre scuole. L'aritmetica, la geometria elementare, la trigonometria e l'algebra elementare erano ormai sostanzialmente complete. Seguì il passaggio alla matematica superiore, alla matematica delle grandezze variabili.

Non è da pensare, tuttavia, che lo sviluppo della matematica elementare sia cessato in questo periodo; ad esempio, nuovi risultati sono stati scoperti e vengono scoperti costantemente oggi nella geometria elementare. Inoltre, è proprio grazie al successivo sviluppo della matematica superiore che ora comprendiamo più chiaramente la natura essenziale della stessa matematica elementare. Ma il ruolo principale in matematica era ora assunto dai concetti di grandezza variabile, funzione e limite. I problemi che hanno portato dalla matematica elementare alla matematica superiore vengono oggi chiariti e risolti dai concetti e dai metodi della matematica superiore (a volte non sono affatto risolvibili con metodi elementari), e ci sono altri problemi che possono essere formulati in termini di matematica elementare ma che servono ancora oggi come fonte di risultati più generali e persino di intere teorie. Ne sono esempi la già citata teoria dei sistemi regolari di figure o problemi di teoria dei numeri elementari nella loro formulazione ma tutt'altro che elementari nei metodi con cui vengono risolti. Per ulteriori dettagli il lettore può consultare il capitolo X.

§ 6. La matematica delle grandezze variabili

1. Variabile e funzione.

Nel XVI secolo lo studio del movimento era il problema centrale della fisica. Le scienze fisiche furono indotte a questo problema, e allo studio di altri problemi riguardanti l'interdipendenza di grandezze variabili, dalle esigenze della vita pratica e dall'intero sviluppo della scienza stessa.

Come riflesso delle proprietà generali del cambiamento, in matematica sorsero i concetti di grandezza variabile e di funzione, e fu questa fondamentale ampliamento dell'oggetto della matematica che determinò il passaggio a una nuova fase, alla matematica delle grandezze variabili.

La legge del moto di un corpo lungo una data traiettoria, ad esempio lungo una linea retta, è definita dal modo in cui la distanza percorsa dal corpo aumenta con il tempo.

Galileo (1564-1642) scoprì così la legge della caduta dei gravi, stabilendo che lo spazio percorso nella caduta cresce proporzionalmente al quadrato del tempo. Questo fatto si esprime nella ben nota formula:

$$s = \frac{gt^2}{2}, \qquad (1)$$

dove g è pari a circa 9,81 m/sec^2.

In generale la legge del moto esprime la distanza percorsa nel tempo t. Qui il tempo t e la distanza s sono rispettivamente la variabile "indipendente" e la "dipendente", e il fatto che ad ogni tempo t corrisponda una definita distanza s è ciò che si intende dicendo che la distanza s è una funzione del tempo t.

I concetti matematici di variabile e funzione sono la generalizzazione astratta di variabili concrete (come tempo, distanza, velocità, angolo di rotazione e area di superficie tracciata) e delle interdipendenze tra di esse (la distanza dipende dal tempo e così via). Come il concetto di numero reale è l'immagine astratta del valore effettivo di una grandezza arbitraria, così la "variabile" è l'immagine astratta di una grandezza variabile, che assume valori diversi durante il processo considerato. Una variabile matematica x è "qualcosa" o, più precisamente, "qualsiasi cosa" che può assumere vari valori numerici. Questo è il significato di una variabile in generale; in particolare, possiamo intendere con esso il tempo, la distanza o qualsiasi altra grandezza variabile.

Allo stesso modo, una funzione è l'immagine astratta della dipendenza di una grandezza da un'altra. L'asserzione che y è una funzione di x significa in matematica solo che ad ogni possibile valore di x corrisponde un determinato valore di y. Questa corrispondenza tra i valori di y e i valori di x è chiamata funzione. Ad esempio, secondo la legge della caduta dei corpi, la distanza percorsa corrisponde al tempo di caduta secondo la formula (1). La distanza è una funzione del tempo. Vediamo alcuni altri esempi.

L'energia di un corpo che cade è espressa dalla sua massa e dalla sua velocità secondo la formula

$$E = \frac{mv^2}{2}. \qquad (2)$$

Per un dato corpo l'energia è funzione della velocità v.

Per una legge familiare la quantità di calore generata in un conduttore nell'unità di tempo dal passaggio di una corrente elettrica è espressa dalla formula

$$Q = \frac{RI^2}{2}, \qquad (3)$$

dove I è l'intensità della corrente e R è la resistenza del conduttore. Per una data resistenza ad ogni corrente I corrisponde una determinata quantità di calore Q, generata nell'unità di tempo. Cioè, Q è una funzione di I.

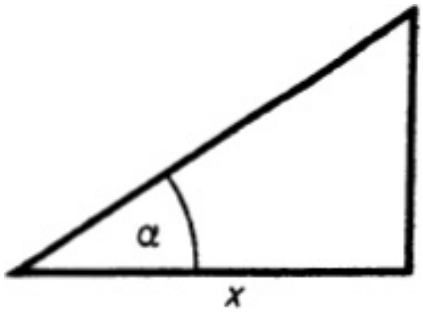


Figura 5

L'area S di un triangolo rettangolo con un dato angolo acuto α e il lato corrispondente x (vedi figura 5) è espressa dalla formula

$$S = \tfrac{1}{2} x^2 \tan \alpha. \qquad (4)$$

Per un dato angolo α l'area è funzione del lato x.

Tutte queste formule (1)-(4) possono essere unite in una

$$y = \tfrac{1}{2} ax^2. \qquad (5)$$

Questa formula generale rappresenta una transizione dalle grandezze variabili concrete t, s, E, Q, v e così via alle variabili generali x e y, e dalle dipendenze concrete (1), (2), (3), (4) alla loro forma generale (5). La meccanica e la teoria dell'elettricità hanno a che fare con le formule concrete (1), (2), (3), che mettono in relazione grandezze concrete, ma la teoria matematica delle funzioni si occupa della formula generale (5), senza associare questa formula ad alcuna grandezza concreta.

Il successivo grado di astrazione dal concreto consiste nell'esaminare non una data dipendenza di y da x, come $y = \frac{1}{2}ax^2$, $y = \sin x$, $y = \log x$ e così via, ma la dipendenza generale di y da x, espressa nella formula astratta

$$y = f(x).$$

Questa formula afferma che la grandezza y è in generale una funzione di x; cioè a ciascun valore assunto da x corrisponde, in un modo o nell'altro, un valore definito y. L'oggetto della matematica consiste quindi non solo di determinate funzioni date ($y = \frac{1}{2}ax^2$, $y = \sin x$ e così via), ma di funzioni arbitrarie (più precisamente, più o meno arbitrarie). Questi gradi di astrazione, prima dalle grandezze concrete e poi dalle funzioni concrete, sono analoghi ai gradi di astrazione osservati nella formazione del concetto di numero intero: in primo luogo, l'astrazione da insiemi concreti di oggetti ha portato al concetto di numeri interi (1, 3, 12 e così via), e poi un'ulteriore astrazione ha portato al concetto di numero intero arbitrario in generale. Questa generalizzazione è il risultato di una profonda interazione tra analisi e sintesi: analisi di interrelazioni separate e sintesi, sotto forma di nuovi concetti, delle loro caratteristiche comuni.

Il ramo della matematica dedicato allo studio delle funzioni è chiamato analisi, o spesso analisi infinitesimale, poiché uno degli elementi più importanti nello studio delle funzioni è il concetto di infinitesimo (il senso e l'importanza di questo concetto sono spiegati nel capitolo II).

Poiché una funzione è l'immagine astratta di una dipendenza di una grandezza da un'altra, possiamo dire che l'analisi ha ad oggetto le dipendenze tra grandezze variabili, non tra una grandezza concreta e l'altra ma tra variabili in generale, astraendo dal loro contenuto. Un'astrazione di questo tipo garantisce una grande ampiezza di applicazione, poiché una formula o un teorema contengono un numero infinito di possibili casi concreti. Un esempio di ciò è già dato dalle nostre semplici formule (1)-(5). Diventa così evidente la completa analogia dell'analisi con l'aritmetica e l'algebra. Tutti hanno origine da determinati problemi pratici e danno un'espressione astratta generale ai rapporti concreti nel mondo reale.

2. Geometria analitica e analisi.

Così il nuovo periodo della matematica, che inizia nel XVII secolo, può essere definito come il periodo della nascita e dello sviluppo dell'analisi. (Questo è il terzo dei grandi periodi menzionati prima.) Naturalmente va inteso che nessuna teoria nasce dalla semplice formazione di nuovi concetti, che l'analisi non potrebbe risultare dalla semplice esistenza dei concetti di variabile e di funzione. Per la fondazione di una teoria, e tanto più di una branca completa della scienza come l'analisi matematica, è necessario che i nuovi concetti diventino, per così dire, attivi, che tra loro si scoprano nuovi rapporti e che permettano la soluzione di nuovi problemi.

Ma soprattutto, nuovi concetti possono nascere e svilupparsi, e diventare più generali e precisi, solo sulla base degli stessi problemi che permettono di risolvere, solo attraverso quei teoremi di cui fanno parte. I concetti di variabile e funzione non sono sorti in forma completa

nella mente di Galileo, Cartesio, Newton o di chiunque altro. Essi vennero in mente a molti matematici (ad esempio a Napier a proposito dei logaritmi) e con Newton e Leibniz assunsero gradualmente una forma più o meno chiara, ma non ancora definitiva, per divenire ancor più precisi e generali nel successivo sviluppo dell'analisi. La loro definizione attuale è stata stabilita solo nel XIX secolo, ma anch'essa non è assolutamente rigorosa o del tutto definitiva. Lo sviluppo del concetto di funzione continua anche oggi.

L'analisi matematica si basava sul materiale fornito dalla nuova scienza della meccanica e su problemi di geometria e di algebra. Il primo passo definitivo verso la matematica delle grandezze variabili fu la comparsa nel 1637 della “geometria” di Cartesio, dove furono poste le basi per la cosiddetta geometria analitica. Le idee di base di Cartesio sono le seguenti.

Supponiamo che ci venga data, ad esempio, l'equazione

$$x^2 + y^2 = a^2. \qquad (6)$$

In algebra x e y erano intese come incognite; e poiché l’equazione data non consente di determinarle, essa non presentava un interesse essenziale per l’algebra. Cartesio, invece, considerò x e y come variabili, non come incognite da ricavare dall’equazione: l’equazione esprimeva così l’interdipendenza di due variabili. Portando tutti i termini al primo membro, un’equazione di questo tipo si può scrivere nella forma generale:

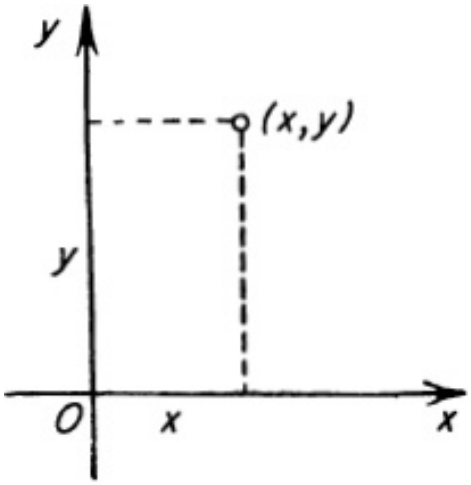


Figura 6.

$$F(x, y) = 0.$$

Cartesio introdusse inoltre nel piano le coordinate x, y, che oggi chiamiamo cartesiane (figura 6). In questo modo, a ogni coppia di valori x e y corrisponde un punto e, viceversa, a ogni punto corrisponde una coppia di coordinate x, y. L’equazione $F(x, y) = 0$ determina quindi il luogo geometrico dei punti del piano le cui coordinate soddisfano l’equazione. In generale, questo luogo sarà una curva. Per esempio, l’equazione (6) determina una circonferenza di raggio a con centro nell’origine. Infatti, come si vede nella figura 7, per il teorema di Pitagora $x^2 + y^2$ è il quadrato della distanza dall’origine O al punto M di coordinate x e y. L’equazione

(6) rappresenta dunque il luogo geometrico dei punti la cui distanza dall'origine è uguale ad *a*, cioè una circonferenza.

Viceversa, un luogo geometrico dei punti, dato da una condizione geometrica, può essere dato anche da un'equazione che esprime la stessa condizione nel linguaggio dell'algebra mediante coordinate. Ad esempio, la condizione geometrica che definisce la circonferenza di un cerchio, cioè che sia un luogo geometrico di punti equidistanti da un punto dato, può essere espressa in linguaggio algebrico dall'equazione (6).

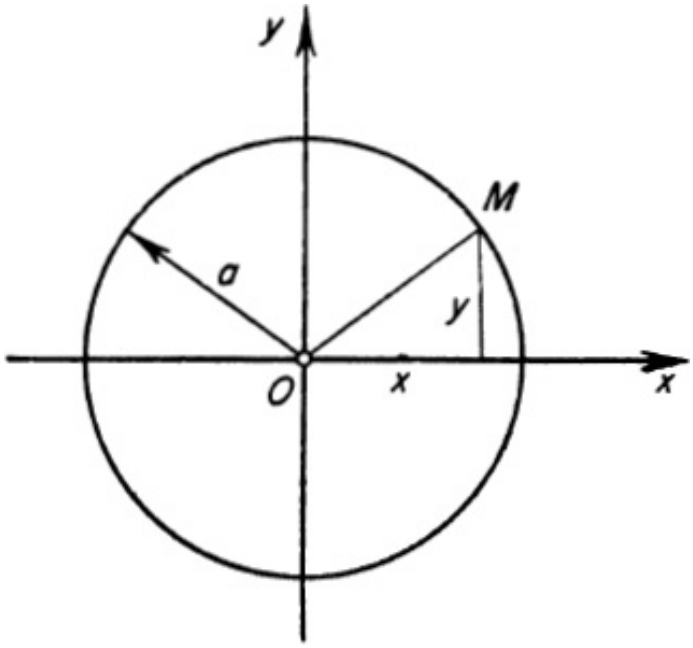


Figura 7.

Quindi il problema generale e il metodo generale della geometria analitica sono i seguenti: rappresentiamo una data equazione in due variabili mediante una curva sul piano, e dalle proprietà algebriche dell'equazione indaghiamo le proprietà geometriche della curva corrispondente; e viceversa, dalle proprietà geometriche della curva troviamo l'equazione, e poi dalle proprietà algebriche dell'equazione investighiamo le proprietà geometriche della curva. In questo modo i problemi geometrici possono essere ridotti ad algebrici, e quindi infine al calcolo.

Il contenuto della geometria analitica sarà discusso in dettaglio nel capitolo III. Vogliamo ora attirare l'attenzione sul fatto che, come risulta evidente dalla nostra breve spiegazione, esso ebbe origine dall'unione della geometria, dell'algebra e dell'idea generale di grandezza variabile. Il principale contenuto geometrico degli inizi della geometria analitica era la teoria delle sezioni coniche, dell'ellisse, dell'iperbole e della parabola. Questa teoria, come abbiamo sottolineato, fu sviluppata dagli antichi greci; i risultati di Apollonio contenevano già in forma geometrica le equazioni delle sezioni coniche. L'unione di questo contenuto geometrico con la forma algebrica, sviluppatasi dopo l'epoca dei Greci, e con l'idea generale di grandezza variabile, derivante dallo studio del movimento, ha prodotto la geometria analitica.

Presso i Greci le sezioni coniche erano un argomento di interesse puramente matematico, ma al tempo di Cartesio erano di importanza pratica per l'astronomia, la meccanica e la tecnologia. Keplero (1571-1630) scoprì che i pianeti si muovono attorno al sole secondo ellissi e Galileo stabilì il fatto che un corpo lanciato in aria, sia esso un sasso o una palla di cannone, si muove lungo una parabola (in prima approssimazione, se si trascura la resistenza dell'aria). Di conseguenza, il calcolo delle varie grandezze riferite alle sezioni coniche divenne una necessità urgente, e fu il metodo di Cartesio a risolvere questo problema. Quindi la via al suo metodo fu preparata dal precedente sviluppo della matematica, e il metodo stesso fu creato dalle insistenti esigenze della scienza e della tecnica.

3. Calcolo differenziale e integrale.

Il successivo passo decisivo nella matematica delle grandezze variabili fu compiuto da Newton e Leibniz durante la seconda metà del XVII secolo, con la fondazione del calcolo differenziale e integrale. Questo fu l'effettivo inizio dell'analisi, poiché l'oggetto di questo calcolo sono le proprietà delle funzioni stesse, distinte dall'oggetto della geometria analitica, che sono le figure geometriche. In realtà Newton e Leibniz portarono a compimento solo un'immensa mole di lavoro preparatorio, condiviso da molti matematici e che risale ai metodi per la determinazione di aree e volumi elaborati dagli antichi greci.

Non esporremo qui i concetti fondamentali del calcolo differenziale e integrale e delle teorie di analisi che ad essi seguirono, poiché ciò verrà fatto nei capitoli speciali dedicati a tali teorie. Vogliamo soltanto richiamare l'attenzione sulle fonti del calcolo infinitesimale, che furono soprattutto i nuovi problemi della meccanica e i vecchi problemi della geometria, questi ultimi consistenti nel tracciare una tangente ad una curva data e nel determinare aree e volumi. Questi problemi geometrici erano già stati studiati dagli antichi (basta citare Archimede), e anche da Keplero, Cavalieri ed altri all'inizio del XVII secolo. Ma l'evento decisivo fu la scoperta della straordinaria relazione tra questi due tipi di problemi e la formulazione di un metodo generale per risolverli; questo fu il risultato di Newton e Leibniz.

Questa relazione, che permette di collegare i problemi della meccanica con quelli della geometria, è stata scoperta grazie alla possibilità, derivante dal metodo delle coordinate, di rappresentare graficamente la dipendenza di una variabile da un'altra, o in altre parole di una funzione. Con l'aiuto di questa rappresentazione grafica ci è facile formulare la relazione, già menzionata, tra i problemi della meccanica e della geometria, che fu all'origine del calcolo differenziale e integrale, e di conseguenza descrivere il contenuto generale di questi due tipi di calcolo.

Il calcolo differenziale è fondamentalmente un metodo per trovare la velocità del movimento quando conosciamo la distanza percorsa in un dato momento. Questo problema è risolto dalla “derivazione”. Si scopre che il problema è del tutto equivalente a quello di tracciare una tangente alla curva che rappresenta la dipendenza della distanza dal tempo. La velocità nell'istante t è uguale alla pendenza della tangente alla curva nel punto corrispondente a t (figura 8).

Il calcolo integrale è sostanzialmente un metodo per trovare la distanza percorsa quando è nota la velocità, o più in generale per trovare il risultato totale dell'azione di grandezza variabile. Questo problema è ovviamente l'inverso del problema del calcolo differenziale (il problema di trovare la velocità); si risolve con l'"integrazione". Risulta che il problema dell'integrazione è del tutto equivalente a quello di trovare l'area sotto la curva che rappresenta la dipendenza della velocità dal tempo. La distanza percorsa nell'intervallo di tempo compreso tra l'istante t_1 e l'istante t_2 è pari all'area sotto la curva compresa tra le rette corrispondenti sul grafico ai valori t_1 e t_2 (figura 9).

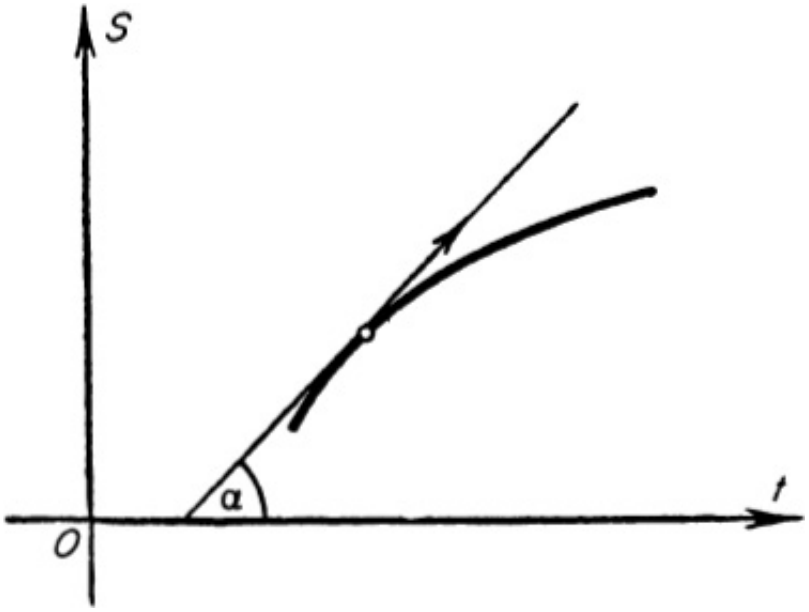


Figura 8.

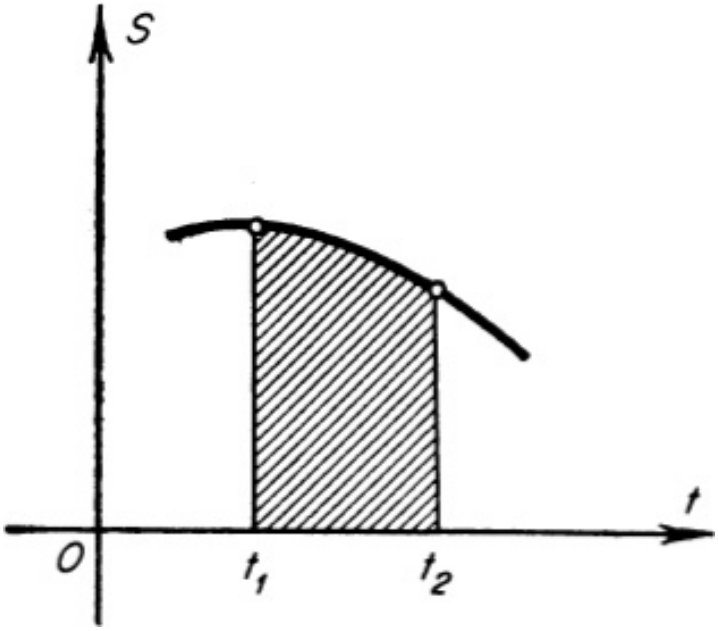


Figura 9.

Astraendo dalla formulazione meccanica dei problemi del calcolo infinitesimale e trattando con le funzioni piuttosto che con la dipendenza della distanza o della velocità dal tempo,

otteniamo il concetto generale dei problemi del calcolo differenziale e integrale in forma astratta.

Fondamentale per il calcolo infinitesimale, come per tutto lo sviluppo successivo dell'analisi, è il concetto di limite, che fu formulato un po' più tardi rispetto agli altri concetti fondamentali di variabile e di funzione. Agli albori dell'analisi il ruolo svolto poi dal limite fu assunto dal concetto alquanto nebuloso di infinitesimo. I metodi per il calcolo vero e proprio della velocità, data la distanza percorsa (cioè derivazione), e della distanza, data la velocità (integrazione), erano fondati sull'unione dell'algebra con il concetto di limite. L'analisi ebbe origine dall'applicazione di questi concetti e metodi ai suddetti problemi di meccanica e geometria (e anche ad alcuni altri problemi; ad esempio, problemi di massimi e minimi). La scienza dell'analisi fu a sua volta assolutamente necessaria per lo sviluppo della meccanica, nella formulazione delle cui leggi i suoi concetti erano già apparsi in forma latente. Ad esempio, la seconda legge di Newton, come formulata dallo stesso Newton, afferma che "la variazione della quantità di moto è proporzionale alla forza agente" (più precisamente: la velocità di variazione della quantità di moto è proporzionale alla forza). Di conseguenza, se vogliamo utilizzare questa legge, dobbiamo essere in grado di definire il tasso di variazione di una variabile, cioè di derivare. (Se enunciamo la legge nella forma che l'accelerazione è proporzionale alla forza, il problema rimane lo stesso, perché l'accelerazione è proporzionale alla velocità di variazione della quantità di moto.) Inoltre, è perfettamente chiaro che per enunciare la legge che governa un movimento quando la forza è variabile (in altre parole, il movimento procede con un'accelerazione variabile), dobbiamo essere in grado di risolvere il problema inverso di trovare una grandezza data la sua velocità di variazione; in altre parole, dobbiamo essere in grado di integrare. Quindi si potrebbe dire che Newton fu semplicemente costretto a inventare la derivazione e l'integrazione per sviluppare la scienza della meccanica.

4. Altri rami dell'analisi.

Insieme al calcolo differenziale e integrale sorsero altri rami dell'analisi: la teoria delle serie (vedi capitolo II, §14), la teoria delle equazioni differenziali (capitoli V e VI) e l'applicazione dell'analisi alla geometria, che in seguito divenne una branca speciale della geometria, chiamata geometria differenziale e che trattava della teoria generale delle curve e delle superfici (capitolo VII). Tutte queste teorie hanno preso vita dai problemi della meccanica, della fisica e della tecnologia.

La teoria delle equazioni differenziali, la branca più importante dell'analisi, ha a che fare con equazioni in cui l'incognita non è più una grandezza ma una funzione, cioè una legge che regola la dipendenza di una grandezza da un'altra o da più altre. È facile capire come siano nate tali equazioni. In meccanica si cerca di determinare l'intera legge del moto di un corpo in determinate condizioni e non solo un valore della velocità o della distanza percorsa. Nella meccanica dei fluidi è necessario trovare la distribuzione della velocità su tutta la massa del fluido in movimento, o in altre parole trovare la dipendenza della velocità da tutte e tre le coordinate spaziali e dal tempo. Analogamente, nella teoria dell'elettricità e del magnetismo dobbiamo trovare l'intensità del campo in tutto lo spazio; cioè la dipendenza di questa intensità dalle stesse tre coordinate spaziali, e analogamente in altri casi.

Problemi di questo genere sorsero continuamente nei vari rami della meccanica, compresa l'idrodinamica e la teoria dell'elasticità, nell'acustica, nella teoria dell'elettricità e del magnetismo e nella teoria del calore. Fin dalla sua nascita l'analisi rimase in stretto contatto con la meccanica e con la fisica in generale, e le sue conquiste più importanti furono invariabilmente legate alla soluzione dei problemi posti dalle scienze esatte. A partire da Newton, i più grandi analisti, D. Bernoulli (1700-1782), L. Euler (1707-1783), J. Lagrange (1736-1813), H. Poincaré (1854-1912), M. V. Ostrogradskij (1801-1861) e A. M. Lyapunov (1857-1918), nonché molti altri, che gettarono nuove basi nell'analisi, partirono di regola dai problemi urgenti della fisica contemporanea.

Sorsero così nuove teorie: in diretto collegamento con la meccanica, Eulero e Lagrange fondarono una nuova branca dell'analisi, chiamata calcolo delle variazioni (vedi capitolo VIII), e alla fine dell'Ottocento Poincaré e Lyapunov, partendo sempre dai problemi della meccanica, fondarono la cosiddetta teoria qualitativa delle equazioni differenziali (vedi capitolo V, §7).

Nel XIX secolo l'analisi si arricchì di un nuovo importante ramo, la teoria delle funzioni di variabile complessa (vedi capitolo IX). I suoi rudimenti si trovano nelle opere di Eulero e di alcuni altri matematici, ma la sua trasformazione in una teoria ben formata avvenne a metà del XIX secolo e fu portata avanti in gran parte dal matematico francese Cauchy (1789-1857). Questa teoria conobbe rapidamente uno sviluppo imponente con numerosi risultati significativi che permisero ai matematici di penetrare più a fondo molte leggi dell'analisi e trovarono importanti applicazioni nei problemi della matematica stessa, della fisica e della tecnologia.

L'analisi si è sviluppata rapidamente; non solo costituì il centro e la parte più importante della matematica, ma penetrò anche nelle regioni più antiche: l'algebra, la geometria e persino la teoria dei numeri. L'algebra cominciò a essere pensata fondamentalmente come la dottrina delle funzioni espresse sotto forma di polinomi di una o più variabili.[23] La geometria analitica e differenziale cominciarono a dominare il campo della geometria. Già da Eulero furono introdotti metodi di analisi nella teoria dei numeri e costituirono così l'inizio della cosiddetta teoria analitica dei numeri, che contiene alcune delle conquiste più profonde della scienza dei numeri interi.

Attraverso l'influenza dell'analisi, con i suoi concetti di variabile, funzione e limite, tutta la matematica fu permeata dall'idea di movimento e cambiamento, e quindi di dialettica. Allo stesso modo, essenzialmente attraverso l'analisi, la matematica è stata influenzata dalle scienze esatte e dalla tecnologia e ha avuto a sua volta un ruolo nel loro sviluppo, poiché è stata il mezzo per dare espressione esatta alle loro leggi e per risolvere i loro problemi. Come presso i Greci la matematica era fondamentalmente geometria, si può dire che dopo Newton fu fondamentalmente analisi. Naturalmente l'analisi non ha assorbito completamente tutta la matematica; nella geometria, nella teoria dei numeri e nell'algebra i problemi e i metodi caratteristici di queste scienze furono ovunque continuati. Così nel XVII secolo nacque, accanto alla geometria analitica, un'altra branca della geometria, la geometria proiettiva, nella quale i metodi puramente geometrici giocavano un ruolo dominante. Ha avuto origine

principalmente nei problemi della rappresentazione di oggetti su un piano (proiezione), e di conseguenza è particolarmente utile nella geometria descrittiva.

Allo stesso tempo si sviluppò una nuova importante branca della matematica, la teoria della probabilità, che prende come oggetto le uniformità osservabili in grandi masse di fenomeni, come una lunga serie di colpi di fucile o di lanci di una moneta. Nel periodo successivo acquistò una particolare importanza nella fisica e nella tecnologia ed il suo sviluppo fu condizionato dai problemi che le provenivano da quei rami della scienza. La caratteristica di questa teoria è che si occupa delle leggi degli “eventi casuali”, fornendo metodi matematici per l’indagine delle irregolarità che necessariamente compaiono negli eventi casuali. Le caratteristiche fondamentali della teoria della probabilità saranno spiegate nel capitolo XI.

5. Applicazioni dell'analisi.

L'analisi in tutti i suoi rami ha fornito alla fisica e alla tecnologia metodi potenti per la soluzione di problemi di vario tipo. Abbiamo già menzionato il primo di questi: trovare il tasso di variazione di una grandezza quando sappiamo come la grandezza stessa dipende dal tempo; trovare l'area delle figure curvilinee e i volumi dei solidi; e trovare il risultato totale di un processo o di un altro o l'azione totale di grandezza variabile. Il calcolo integrale permette quindi di determinare il lavoro compiuto da un gas in espansione al variare della pressione secondo una legge ben nota; lo stesso calcolo integrale ci permette di calcolare, ad esempio, l’intensità di un campo elettrico con un sistema di cariche dato arbitrariamente, basando il nostro lavoro sulla legge di Coulomb che determina l’intensità di un campo risultante da una carica puntiforme, e così via.

Inoltre, l'analisi ha fornito un metodo per trovare i valori massimo e minimo di una grandezza in determinate condizioni. Quindi con l'aiuto dell'analisi è facile determinare la forma di una cisterna cilindrica che a parità di volume avrà la superficie più piccola e di conseguenza richiederà il minor dispendio di materiale. Risulta che la cisterna avrà questa proprietà se la sua altezza sarà pari al diametro della sua base (figura 10). L'analisi permette di determinare la forma della curva lungo la quale un corpo deve rotolare per cadere nel più breve tempo possibile da un dato punto a un altro (questa curva è la cosiddetta cicloide; figura 11).

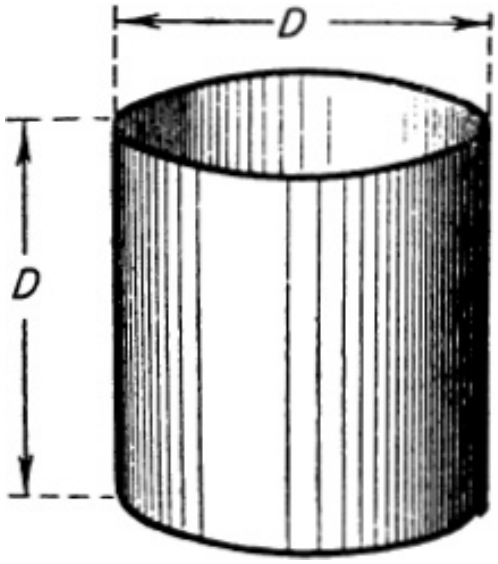


Figura 10.

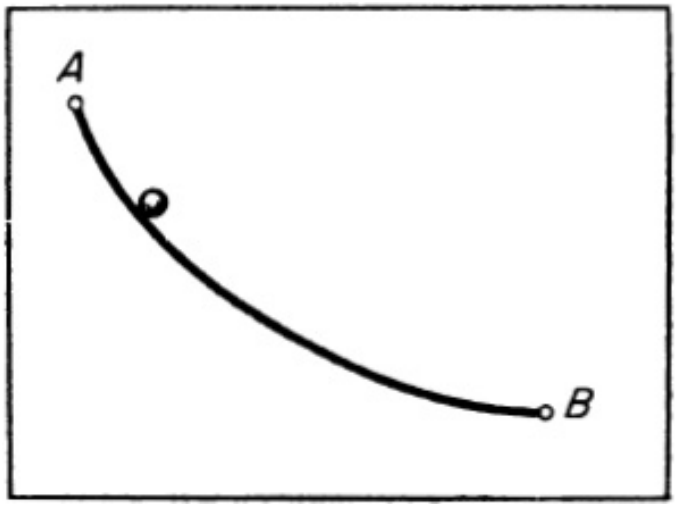


Figura 11.

Per la soluzione di questi ed altri problemi il lettore può rivolgersi ai capitoli II e VIII.

L'analisi, o più precisamente la teoria delle equazioni differenziali, ci consente non solo di trovare valori separati per grandezze variabili ma anche di determinare funzioni incognite; cioè trovare leggi di dipendenza di alcune variabili da altre. Abbiamo così la possibilità, sulla base delle leggi generali dell'elettricità, di calcolare come varia la corrente nel tempo in un circuito con resistenza, capacità e autoinduzione arbitrarie. Possiamo determinare le leggi per la distribuzione delle velocità nell'intera massa di un fluido in determinate condizioni. Possiamo dedurre leggi generali per la vibrazione delle corde e delle membrane, e per la propagazione delle vibrazioni nei vari mezzi; qui ci riferiamo alle onde sonore, alle onde elettromagnetiche o alle vibrazioni elastiche propagate attraverso la Terra da terremoti o esplosioni. Tra parentesi possiamo osservare che in tal modo vengono forniti nuovi metodi per la ricerca di minerali utili e per l'esecuzione di indagini molto al di sotto della superficie della Terra. Singoli problemi di questo tipo si troveranno nei capitoli V e VI.

Infine, l'analisi non solo ci fornisce metodi per risolvere problemi particolari; ci fornisce anche metodi generali per la formulazione matematica delle leggi quantitative delle scienze esatte. Come accennato in precedenza, le leggi generali della meccanica non potrebbero essere formulate matematicamente senza il ricorso ai concetti di analisi, e senza tale formulazione non saremmo in grado di risolvere i problemi della meccanica. Allo stesso modo le leggi generali sulla conduzione del calore, sulla diffusione attraverso materiali porosi, sulla propagazione delle vibrazioni, sul decorso delle reazioni chimiche, sulle leggi fondamentali dell'elettromagnetismo e su molte altre leggi semplicemente non potrebbero essere formulate matematicamente senza i concetti di analisi. Solo come risultato di una tale formulazione queste leggi possono essere applicate ai casi concreti più diversi, fornendo una base per conclusioni matematiche esatte nei problemi speciali della conduzione del calore, delle vibrazioni, delle soluzioni chimiche, dei campi elettromagnetici e di altri problemi della meccanica, dell'astronomia e di tutti i numerosi rami della fisica, della chimica, dell'ingegneria termica, dell'energia, della costruzione di macchine, dell'ingegneria elettrica, ecc.

6. Esame critico dei fondamenti dell'analisi.

Come nella storia della geometria greca la presentazione rigorosa e sistematica data da Euclide portò a compimento un lungo sviluppo precedente, così nello sviluppo dell'analisi si presentò la necessità di porla su una base più solida di quella fornita dai primi creatori dei suoi potenti metodi: Newton, Eulero, Lagrange ed altri. Man mano che l'analisi da loro fondata si estendeva, essa cominciò, da un lato, ad affrontare problemi più profondi e difficili, e dall'altro a richiedere per la sua stessa estensione una base più sistematica e attentamente ragionata. La crescita della teoria ha reso necessaria una sistematizzazione e un'analisi critica dei suoi fondamenti. Per porre una teoria su un solido fondamento è necessario esaminare il suo intero sviluppo e non deve assolutamente essere considerato come un punto di partenza per la teoria stessa, poiché senza la teoria semplicemente non avremmo idea di cosa dobbiamo fondare. Del resto, certi formalisti contemporanei dimenticano questo fatto quando ritengono opportuno fondare e sviluppare una teoria a partire da assiomi che non sono stati selezionati sulla base di alcuna analisi del materiale concreto che dovrebbero riassumere. Ma gli assiomi stessi richiedono una giustificazione del loro contenuto; riassumono solo altro materiale e forniscono una base per la costruzione logica di una teoria.[24]

Il periodo necessario di critica, di sistematizzazione e di fondazione si è verificato nell'analisi all'inizio del secolo scorso. Grazie agli sforzi di numerosi eminenti scienziati questo importante e difficile lavoro è stato portato a termine con successo. In particolare, sono state date definizioni precise per i concetti base di numero reale, variabile, funzione, limite e continuità.

Tuttavia, come abbiamo già avuto modo di ricordare, nessuna di queste definizioni può essere considerata assolutamente rigorosa o definitiva. Lo sviluppo di questi concetti continua. Euclide e tutti i matematici nel corso di 2000 anni dopo di lui considerarono senza dubbio i suoi “Elementi” come il limite pratico del rigore logico. Ma ad una visione contemporanea i fondamenti euclidei della geometria sembrano piuttosto superficiali. Questo esempio storico

mostra che non dovremmo illuderci con alcuna idea di rigore “assoluto” o “finale” nella matematica contemporanea. In una scienza che non sia ancora morta e mummificata, non c'è e non può esserci nulla di perfetto. Ma possiamo dire con sicurezza che i fondamenti dell'analisi quali esistono attualmente corrispondono in modo del tutto soddisfacente ai problemi contemporanei della scienza e alla concezione contemporanea della precisione logica; e in secondo luogo, che il continuo approfondimento di questi concetti e le discussioni che si stanno svolgendo attualmente su di essi non ci danno motivo, e non ci daranno motivo, semplicemente per rifiutarli; queste discussioni ci condurranno a una comprensione nuova, più precisa e più profonda, i cui risultati sono ancora difficili da stimare.

Sebbene la fissazione dei principi fondamentali di una teoria costituisca una sintesi del suo sviluppo, non rappresenta la fine della teoria; al contrario, è favorevole a un ulteriore sviluppo. Questo è esattamente quello che è successo in analisi. In connessione con l'approfondimento dei suoi fondamenti sorse una nuova teoria matematica, creata dal matematico tedesco Cantor negli anni settanta del secolo scorso, vale a dire la teoria generale degli insiemi infiniti di oggetti astratti arbitrari, siano essi numeri, punti, funzioni o qualsiasi altro "elemento". Sulla base di queste idee è nato un nuovo capitolo dell'analisi, la cosiddetta teoria delle funzioni di variabile reale, i cui concetti, insieme a quelli dei fondamenti dell'analisi e della teoria degli insiemi, sono spiegati nel capitolo XV. Allo stesso tempo le idee generali della teoria degli insiemi penetrarono in ogni ramo della matematica. Ma questo “punto di vista della teoria degli insiemi” è inseparabilmente connesso con una nuova fase nello sviluppo della matematica, che ora considereremo brevemente.

§ 7. La matematica contemporanea

1. Il carattere più avanzato della matematica odierna.

Ai quattro stadi dello sviluppo della matematica menzionati nel §5 corrispondono naturalmente stadi della nostra educazione matematica, poiché il materiale appreso in ogni stadio del nostro studio consiste, con un discreto grado di approssimazione, del contenuto di base del periodo corrispondente nella storia della matematica.

I risultati fondamentali dell'aritmetica e della geometria, ottenuti nel primo periodo dello sviluppo della matematica, costituiscono oggetto dell'istruzione primaria e sono noti a tutti noi. Quando, ad esempio, determiniamo la quantità di materiale necessario per realizzare un certo lavoro, diciamo per rivestire un pavimento, stiamo già utilizzando questi primi risultati della matematica. Le conquiste più importanti del secondo periodo, quello della matematica elementare, vengono insegnate nelle scuole superiori. I risultati fondamentali del terzo periodo, i fondamenti dell'analisi, la teoria delle equazioni differenziali, l'algebra superiore e così via, costituiscono l'istruzione matematica di un ingegnere; vengono studiati in tutte le scuole di insegnamento superiore, tranne quelle prettamente umanistiche. In questo modo le idee fondamentali e i risultati della matematica di quel periodo sono ampiamente conosciuti e se ne servono, in una certa misura, quasi tutti gli ingegneri e gli scienziati.

D'altra parte, le idee e i risultati dell'attuale periodo della matematica sono studiati quasi esclusivamente nei dipartimenti universitari di matematica e fisica. Oltre agli specialisti in matematica, vengono utilizzati dai ricercatori nei campi della meccanica e della fisica, nonché in alcuni dei rami più recenti della tecnologia. Naturalmente, questo non significa affatto che non abbiano applicazione pratica, ma poiché rappresentano i risultati più recenti della scienza, sono naturalmente più complicati. Di conseguenza, quando passiamo ora a una descrizione generale dell'ultimo stadio dello sviluppo della matematica, non possiamo più ritenere che tutto ciò cui accenneremo brevemente sia del tutto chiaro. Cercheremo di presentare in poche righe il carattere più generale delle nuove branche della matematica; il loro contenuto verrà spiegato più dettagliatamente nei capitoli corrispondenti del libro.

Se la presente sezione sembra eccessivamente difficile, può essere tralasciata in prima lettura e ripresa dopo lo studio dei capitoli speciali.

2. Geometria.

L'inizio dello sviluppo attuale della matematica è caratterizzato da profondi cambiamenti in tutti i suoi campi fondamentali: algebra, geometria e analisi. Questo cambiamento può forse essere seguito più chiaramente nel campo della geometria. Nel 1826 Lobačevskij, e quasi contemporaneamente a lui il matematico ungherese Janos Bolyai, sviluppò la nuova geometria non euclidea. Le idee di Lobačevskij erano lungi dall'essere immediatamente comprese da tutti i matematici. Erano troppo audaci e inaspettate. Ma da questo momento cominciò un nuovo sviluppo fondamentale della geometria; la concezione stessa di cosa si intende per geometria venne cambiata. Il suo oggetto e la gamma delle sue applicazioni furono rapidamente ampliati. Il passo più importante in questa direzione, dopo Lobačevskij, fu compiuto nel 1854 dal celebre matematico tedesco Riemann. Formulò chiaramente l'idea generale che un numero illimitato di “spazi” potevano essere indagati dalla geometria, e allo stesso tempo ne indicò il possibile significato nel mondo reale. Nel nuovo sviluppo della geometria emergevano due caratteristiche fondamentali.

In primo luogo la geometria precedente studiava solo le forme e i rapporti spaziali del mondo materiale, e poi solo nella misura in cui compaiono nel quadro della geometria euclidea, ma ora l'oggetto della geometria cominciò ad includere anche molte altre forme e rapporti del mondo reale, purché solo fossero simili a quelli spaziali e quindi consentissero l'uso di metodi geometrici. Il termine “spazio” assunse così in matematica un significato nuovo, più ampio e allo stesso tempo più speciale. Allo stesso tempo, i metodi della geometria divennero molto più ricchi e vari. A loro volta ci forniscono mezzi più completi per conoscere il mondo fisico che ci circonda, il mondo da cui è stata astratta la geometria nella sua forma originale.

In secondo luogo, anche nella geometria euclidea furono compiuti importanti progressi: in essa furono studiate le proprietà di figure incomparabilmente più complicate, comprendenti anche insiemi arbitrari di punti. Emerse anche un atteggiamento fondamentalmente nuovo nei confronti delle proprietà delle figure oggetto dell'indagine. Furono distinti gruppi separati di proprietà, che potevano essere studiati in astrazione dagli altri, e proprio questa astrazione all'interno della geometria diede origine a molti rami caratteristici della materia, che essenzialmente divennero "geometrie" indipendenti. Lo sviluppo della geometria in tutte queste direzioni continua e vengono studiati sempre più nuovi “spazi” e le loro “geometrie”: lo spazio di Lobačevskij, lo spazio proiettivo, lo spazio euclideo e altri spazi di varie dimensioni, in particolare lo spazio quadridimensionale, gli spazi di Riemann, gli spazi di Finsler, gli spazi topologici, ecc. Queste teorie trovano importante applicazione nella stessa matematica, al di fuori della geometria, e anche nella fisica e nella meccanica; particolarmente degne di nota sono le loro applicazioni nella teoria della relatività della fisica contemporanea, che è una teoria dello spazio, del tempo e della gravitazione. Da quanto detto è chiaro che si tratta qui di un cambiamento qualitativo della geometria.

Le idee della geometria contemporanea e alcuni elementi della teoria dei vari spazi in essa indagati saranno spiegati nei capitoli XVII e XVIII.

3. Algebra.

Anche l'algebra ha subito un cambiamento qualitativo. Nella prima metà del XIX secolo sorsero nuove teorie che ne modificarono il carattere, ne ampliarono l'oggetto e il campo di applicazione.

Nella sua forma originaria, come sottolineato nel §5, l'algebra si occupava di operazioni matematiche sui numeri considerati da un punto di vista formale, in astrazione da numeri concreti dati. Questa astrazione ha trovato espressione nel fatto che in algebra le grandezze sono indicate da lettere, sulle quali vengono eseguiti calcoli secondo regole formali ben note.

L'algebra contemporanea conserva questa base ma la amplia in misura considerevole. Ora considera le "grandezze" di natura molto più generale dei numeri, e studia le operazioni su queste "grandezze" che sono in una certa misura analoghe nelle loro proprietà formali alle operazioni ordinarie dell'aritmetica: addizione, sottrazione, moltiplicazione e divisione. Un esempio molto semplice è offerto dalle grandezze vettoriali, che possono essere sommate con la nota regola del parallelogramma. Ma la generalizzazione operata nell'algebra contemporanea è tale che anche il termine stesso "grandezza" spesso perde il suo significato e si parla più in generale di "elementi" sui quali è possibile compiere operazioni simili a quelle consuete algebriche. Ad esempio, due movimenti compiuti uno dopo l'altro equivalgono evidentemente ad un certo movimento unico, che è la somma dei due; due trasformazioni algebriche di una formula possono equivalere a un'unica trasformazione che produce lo stesso risultato, e così via; e quindi si può parlare di una caratteristica "addizione" di movimenti o trasformazioni. Tutto questo e molto altro ancora viene studiato in forma generale astratta nell'algebra contemporanea.

Le nuove teorie algebriche in questa direzione sorsero nella prima metà del XIX secolo nelle ricerche di numerosi matematici, tra i quali va menzionato in particolare il matematico francese Galois (1811-1832). I concetti, i metodi e i risultati dell'algebra contemporanea trovano importanti applicazioni nell'analisi, nella geometria, nella fisica e nella cristallografia. In particolare, la teoria menzionata alla fine del § 3 riguardante la simmetria dei cristalli, sviluppata da E. S. Fedorov, si basa sull'unione della geometria con una delle nuove teorie algebriche, la cosiddetta teoria dei gruppi.

Come si vede, si tratta qui di una generalizzazione fondamentale e qualitativa dell'argomento dell'algebra con un cambiamento nel concetto stesso di cosa sia l'algebra. Le idee dell'algebra contemporanea e gli elementi base di alcune delle sue teorie saranno spiegati nei capitoli XX e XVI.

4. Analisi.

Anche l'analisi in tutti i suoi rami fece profondi progressi. In primo luogo, come già accennato nel paragrafo precedente, se ne precisarono i fondamenti; in particolare, ai suoi concetti fondamentali furono date definizioni esatte e generali: concetti come funzione, limite, integrale e, infine, il concetto base di grandezza variabile (per il numero reale fu data una definizione rigorosa). Un inizio del processo di porre l'analisi su basi più precise fu fatto dal matematico ceco Bolzano (1781-1848), dal matematico francese Cauchy (1789-1857) e da molti altri. Questa maggiore precisione fu acquisita contemporaneamente ai nuovi sviluppi dell'algebra e della geometria; fu portato a termine nella sua forma oggi ben nota negli anni ottanta del XIX secolo dai matematici tedeschi Weierstrass, Dedekind e Cantor. Come accennato alla fine del §6, Cantor pose anche le basi per la teoria degli insiemi transfiniti, che gioca un ruolo così importante nello sviluppo delle nuove idee in matematica.

L'aumento della precisione nei concetti di variabile e funzione in connessione con la teoria degli insiemi gettò le basi per un ulteriore sviluppo dell'analisi. Si è passati allo studio di funzioni più generali; e nella stessa direzione venne generalizzato anche l'apparato di analisi, cioè il calcolo integrale e differenziale. Sorse così, alle soglie del nostro secolo, la nuova branca dell'analisi già menzionata nel §6, la cosiddetta teoria delle funzioni di variabile reale. Lo sviluppo di questa teoria è legato soprattutto ai matematici francesi Borel, Lebesgue ed altri, nonché a N. N. Luzin (1883-1950) e alla sua scuola. In generale, i rami più recenti dell'analisi sono chiamati analisi moderna in contrapposizione alla precedente analisi cosiddetta classica.

Altre nuove teorie sono emerse nell'analisi. Si è così formato un ramo speciale della teoria dell'approssimazione delle funzioni, che studia le questioni relative alla migliore rappresentazione approssimata di funzioni generali mediante varie funzioni "semplici", soprattutto mediante polinomi, cioè mediante funzioni della forma

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n .$$

La teoria dell'approssimazione delle funzioni ha una grande importanza, se non altro perché pone basi generali per il calcolo pratico delle funzioni, per la sostituzione approssimativa di funzioni complicate con funzioni più semplici. I rudimenti di questa teoria risalgono agli inizi dell'analisi. Il suo orientamento moderno gli è stato dato dal grande matematico russo P. L. Čebyšëv (1821-1894). Questa direzione fu successivamente sviluppata nella cosiddetta teoria costruttiva delle funzioni, soprattutto nei lavori dei matematici sovietici, in particolare S. N. Bernštejn (nato nel 1880), a cui appartengono i risultati più importanti in questo campo. Il capitolo XII tratta dell'approssimazione delle funzioni.

Abbiamo parlato prima dello sviluppo della teoria delle funzioni di variabile complessa. Dobbiamo ancora menzionare la cosiddetta teoria qualitativa delle equazioni differenziali, che ha origine nei lavori di Poincaré (1854-1912) e A. M. Lyapunov (1857-1918), sulla quale si daranno alcune idee nel capitolo V, e anche la teoria delle equazioni integrali. Queste teorie hanno una grande importanza pratica in meccanica, fisica e tecnologia. Pertanto, la teoria qualitativa delle equazioni differenziali fornisce soluzioni a problemi riguardanti la stabilità del movimento e l'azione di meccanismi o di sistemi elettrici vibranti e simili. Stabilità di un

processo significa nel senso più generale che se vengono apportate piccole modifiche ai dati iniziali o alle condizioni del movimento, allora il movimento stesso durante tutto il suo corso cambierà solo leggermente. Non è necessario sottolineare il significato tecnico di questioni di questo genere.

5. Analisi funzionale.

Sul terreno preparato dallo sviluppo dell'analisi e della fisica matematica, insieme alle nuove idee della geometria e dell'algebra, è cresciuta una nuova vasta divisione della matematica, la cosiddetta analisi funzionale, che svolge un ruolo eccezionalmente importante nella matematica moderna. Molti matematici hanno partecipato alla sua creazione; ricordiamo, ad esempio, il più grande matematico tedesco degli ultimi tempi, Hilbert (1862-1943), il matematico ungherese Riesz (1880-1956) e il matematico polacco Banach (1892-1945). Il capitolo XIX è dedicato all'analisi funzionale.

L'essenza di questa nuova branca della matematica consiste brevemente in quanto segue. Nell'analisi classica la variabile è una grandezza, o "numero", ma nell'analisi funzionale la funzione stessa è considerata la variabile. Le proprietà della funzione data qui non sono determinate in sé ma in relazione ad altre funzioni. Ciò che si studia non è una funzione separata ma un insieme di funzioni caratterizzate da una proprietà o da un'altra; ad esempio, la raccolta di tutte le funzioni continue. Un tale insieme di funzioni costituisce il cosiddetto spazio funzionale. Questo procedimento corrisponde, ad esempio, al fatto che possiamo considerare la raccolta di tutte le curve su una superficie o di tutti i possibili movimenti di un dato sistema meccanico, definendo così le proprietà delle singole curve o movimenti nella loro relazione con altre curve o movimenti.

Il passaggio dallo studio di funzioni separate a una funzione variabile è simile al passaggio dai numeri incogniti x, y alle variabili x, y; cioè è simile all'idea di Cartesio menzionata in un paragrafo precedente. Sulla base di questa idea Cartesio ha prodotto la sua famosa unione di algebra e geometria, di un'equazione e di una curva, che è uno degli elementi più importanti nello sviluppo dell'analisi. Allo stesso modo, l'unione del concetto di funzione variabile con le idee dell'algebra e della geometria contemporanee ha prodotto la nuova analisi funzionale. Come l'analisi fu necessaria per lo sviluppo della meccanica dell'epoca, così l'analisi funzionale fornì nuovi metodi per la soluzione degli attuali problemi della fisica matematica e produsse l'apparato matematico per la nuova meccanica quantistica dell'atomo. La storia si ripete come al solito, ma in modo nuovo, su un piano più alto. Come abbiamo detto, l'analisi funzionale unisce le idee e i metodi fondamentali dell'analisi, dell'algebra moderna e della geometria e, a sua volta, esercita un'influenza sullo sviluppo di questi rami della matematica. I problemi che sorgono nell'analisi classica trovano ora soluzioni nuove e più generali, spesso quasi in un unico passaggio, mediante l'analisi funzionale. Qui, come in un fuoco, sono raccolte insieme, in modo molto produttivo, le idee più generali e astratte della matematica moderna.

Da questo breve abbozzo, da questa semplice enumerazione dei nuovi orientamenti dell'analisi (teoria delle funzioni di variabile reale, teoria dell'approssimazione delle funzioni, teoria qualitativa delle equazioni differenziali, teoria delle equazioni integrali e analisi

funzionale) si vede che qui si tratta in realtà di una fase essenzialmente nuova nello sviluppo dell'analisi.

6. Matematica computazionale e logica matematica.

In tutti i periodi il livello tecnico dei mezzi di calcolo ha avuto un'influenza essenziale sui metodi matematici. Ma i mezzi di calcolo di cui disponiamo fino ai tempi più recenti erano molto limitati. Gli apparecchi più semplici, come l'abaco, le tavole dei logaritmi e il regolo calcolatore, la macchina calcolatrice, ed infine i calcolatori più complicati e la macchina calcolatrice automatica, furono gli strumenti fondamentali per il calcolo esistenti fino agli anni Quaranta del XX secolo. Questi strumenti permettevano di eseguire più o meno velocemente le singole operazioni di addizione, moltiplicazione, ecc. Ma per portare al risultato numerico finale i problemi pratici che si pongono oggigiorno è necessario un numero colossale di tali operazioni, che si susseguono in un programma complicato che a volte dipende dai risultati ottenuti nel corso del calcolo. La soluzione di tali problemi si è rivelata praticamente impossibile o del tutto priva di valore a causa della lunghezza del processo di soluzione. Ma negli ultimi dieci anni si è verificato un cambiamento radicale nell'intera scienza dei calcoli. Le moderne macchine calcolatrici, costruite secondo nuovi principi, ci permettono di effettuare calcoli con una velocità eccezionalmente grande e allo stesso tempo di eseguire complicate catene di calcoli automaticamente, secondo programmi estremamente flessibili predisposti in anticipo. Alcune delle questioni legate alla costruzione e all'importanza delle moderne macchine calcolatrici saranno discusse nel capitolo XIV.

Le nuove tecniche non solo ci permettono di effettuare indagini che prima erano del tutto impraticabili, ma ci portano anche a modificare la nostra stima del valore di molti risultati matematici ben noti. Ad esempio, hanno dato uno stimolo speciale allo sviluppo di metodi approssimativi; cioè metodi che permettono, mediante una catena di operazioni elementari, di raggiungere un risultato numerico desiderato con una precisione sufficientemente grande. I metodi matematici stessi devono ora essere valutati dal punto di vista della loro idoneità per le macchine corrispondenti.

La logica matematica è strettamente connessa con lo sviluppo delle tecniche di calcolo. Si è sviluppata soprattutto a partire da difficoltà interne alla matematica stessa e ha per oggetto l'analisi della dimostrazione matematica. È essa stessa un ramo della matematica e comprende quelle parti della logica generale che possono essere formulate e sviluppate oggettivamente mediante il metodo matematico.

Sebbene da un lato la logica matematica risalga così alle fonti e ai fondamenti stessi della matematica, dall'altro è strettamente connessa con le più moderne questioni della tecnica computazionale. Naturalmente, ad esempio, una dimostrazione che porta alla messa a punto di un processo definito e preassegnato, permettendoci di avvicinarci ad un risultato desiderato con un grado arbitrario di accuratezza, è essenzialmente diversa dalle prove più astratte sull'esistenza del risultato dato.

Anche qui si pone una gamma caratteristica di domande riguardanti il grado di generalità possibile nei problemi che possono essere affrontati con un metodo completamente definito in

anticipo in ogni fase. In questo senso sono stati raggiunti risultati profondi in logica matematica, risultati estremamente importanti dal punto di vista epistemologico generale.

Non sarebbe esagerato affermare che con lo sviluppo delle nuove tecniche di calcolo e le conquiste della logica matematica è iniziato un nuovo periodo nella matematica moderna, caratterizzato dal fatto che il suo oggetto non è solo lo studio di questo o quell'oggetto ma anche tutti i modi e i mezzi con cui tale oggetto può essere definito; non solo alcuni problemi, ma anche tutti i metodi possibili per risolverli.

A quanto detto è solo necessario aggiungere che anche nei rami più antichi della matematica, la teoria dei numeri, la geometria euclidea, l'algebra classica e l'analisi, e la teoria della probabilità, il rapido sviluppo è continuato durante tutto il periodo della matematica moderna, così che questi campi sono stati arricchiti di molte nuove idee e risultati fondamentali; ricordiamo, ad esempio, i risultati raggiunti nella teoria dei numeri e nella geometria dello spazio quotidiano dai matematici russi e sovietici P. L. Čebyšëv, E. S. Fedorov, I. M. Vinogradov ed altri. Lo sviluppo su un ampio fronte della teoria della probabilità è stato collegato alle regolarità straordinariamente importanti osservabili nella fisica statistica e nei problemi contemporanei della tecnologia.

7. Aspetti caratteristici della matematica moderna.

Quali sono le caratteristiche più generali della matematica moderna nel suo insieme, che la distinguono dai precedenti sviluppi della geometria, dell’algebra e dell’analisi?

Innanzitutto l'immensa estensione della materia della matematica e delle sue applicazioni. Una tale estensione della materia e del campo di applicazione rappresenta un'enorme crescita quantitativa e qualitativa, determinata dalla comparsa di nuove potenti teorie e metodi che ci permettono di risolvere problemi fino ad oggi completamente inaccessibili. Questo ampliamento dell'oggetto della matematica è caratterizzato dal fatto che la matematica contemporanea si pone consapevolmente il compito di studiare tutti i possibili tipi di relazioni quantitative e forme spaziali.

Una seconda caratteristica della matematica moderna è la formazione di concetti generali a un nuovo e più elevato livello di astrazione. È proprio questa caratteristica che garantisce la preservazione dell'unità della matematica, nonostante la sua immensa crescita in rami molto diversi. Anche in parti della matematica estremamente distanti tra loro, le somiglianze strutturali vengono messe in luce dai concetti generali e dalle teorie dei nostri giorni. Garantiscono che i metodi matematici contemporanei avranno grande generalità e respiro di applicazione; in particolare, producono una profonda compenetrazione dei rami fondamentali della matematica: geometria, algebra e analisi.

Tra gli aspetti caratteristici della matematica moderna va menzionata anche l'evidente predominanza del punto di vista della teoria degli insiemi. Naturalmente questo punto di vista deve la sua importanza al fatto che riassume in un certo senso il ricco contenuto di tutti i precedenti sviluppi della matematica. Infine, uno dei tratti più caratteristici della matematica moderna è l'analisi profonda dei suoi fondamenti, dell'influenza reciproca dei suoi concetti, della struttura delle sue teorie separate e dei metodi di dimostrazione matematica. Senza una

tale analisi dei fondamenti non sarebbe possibile migliorare o sviluppare ulteriormente i principi e le teorie che hanno portato alle attuali generalizzazioni.

Si può dire che la caratteristica della matematica moderna sia che il suo oggetto non consiste solo di relazioni e forme quantitative date, ma di tutte quelle possibili. In geometria non parliamo solo di rapporti e forme spaziali, ma di tutte le possibili forme simili a quelle spaziali. Nell'algebra parliamo di vari sistemi astratti di oggetti con tutte le possibili leggi delle operazioni su di essi. Nell'analisi non vengono considerate variabili solo le grandezze, ma le funzioni stesse. In uno spazio funzionale sono riunite tutte le funzioni di un dato tipo (tutte le possibili interdipendenze tra le variabili). Riassumendo, si può dire che mentre la matematica elementare si occupa di grandezze costanti, e il periodo successivo di grandezze variabili, la matematica contemporanea è la matematica di tutte le possibili (in generale variabili) relazioni quantitative e interdipendenze tra grandezze. Questa definizione è, ovviamente, incompleta, ma sottolinea il tratto caratteristico della matematica moderna che la distingue da quella delle epoche precedenti.[25]

Letture consigliate

Osservazione preliminare dell'edizione inglese. Il testo originale russo di *Mathematics: Its Content, Methods, and Meaning* contiene un elenco di letture consigliate alla fine di ciascuno dei venti capitoli. Nella traduzione inglese tali libri sono stati mantenuti soltanto quando disponibili in inglese. In compenso, gli elenchi includono numerose altre opere facilmente reperibili in quella lingua.

Opere sulla matematica in generale

E. T. Bell, The development of mathematics, 2ª ed., McGraw-Hill, New York, 1945.

R. Courant e H. Robbins, What is mathematics ? Oxford University Press, New York, 1941.

H. Eves e C. V. Newsom, An introduction to the foundations e fundamental concepts of mathematics, Rinehart, New York, 1958.

G. H. Hardy, A mathematician's apology, Macmillan, New York, 1940.

R. L. Wilder, Introduction to the foundations of mathematics, Wiley, New York, 1952.

Opere di carattere storico

R. C. Archibald, Outline of the history of mathematics, 5ª ed., Mathematical Association of America, Oberlin, Ohio, 1941.

F. Cajori, History of mathematics, 2ª ed., Macmillan, New York, 1919.

Euclid, The thirteen books of Euclid's Elements translated with an introduction e commentary by T. L. Heath, 2ª ed., 3 vols., Dover, New York, 1956.

O. E. Neugebauer, The exact sciences in antiquity, Princeton University Press, Princeton, N. J., 1952.

D. E. Smith, History of mathematics, Vol. I. General survey of the history of elementary mathematics, Vol. II. Special topics of elementary mathematics, Dover, New York, 1958.

D. J. Struik, A concise history of mathematics, Dover, New York, 1948.

B. L. van der Waerden, Science awakening, P. Noordhoff, Groningen, 1954.

Note al capitolo I

[1] Qui ci limitiamo a segnalare questo esempio senza ulteriori spiegazioni, per le quali il lettore può rivolgersi al Capitolo XVII.

[2] Infatti ogni collezione di oggetti, sia essa un gregge di pecore o una catasta di legna da ardere, esiste e viene immediatamente percepita in tutta la sua concretezza e complessità. La distinzione in essa di proprietà e rapporti separati è il risultato di un'analisi cosciente. Il pensiero primitivo non fa ancora questa analisi, ma considera l'oggetto solo nel suo insieme. Allo stesso modo, un uomo che non ha studiato musica percepisce una composizione musicale senza distinguere in essa i dettagli della melodia, della tonalità, ecc., mentre allo stesso tempo un musicista analizza facilmente anche una sinfonia complicata.

[3] Nella formazione dei concetti sulle proprietà degli oggetti, come il colore o la numerosità di una collezione, è possibile distinguere tre fasi, che non dobbiamo, ovviamente, cercare di separare troppo nettamente l'una dall'altra. Nella prima fase la proprietà è definita dal confronto diretto degli oggetti: come un corvo, tanti quanti su una mano. Al secondo compare un aggettivo: una pietra nera oppure (l'aggettivo numerico è del tutto analogo) cinque alberi. Nella terza fase la proprietà viene astratta dagli oggetti e può apparire "come tale"; per esempio "nerezza" o il numero astratto "cinque".

[4] La parola «aritmetica», che significa «arte del calcolo», deriva dall'aggettivo greco *arithmētik*□, formato dal sostantivo *arithmós*, «numero». L'aggettivo si riferisce al sostantivo sottinteso *téchnē*, «arte», «tecnica».

[5] Ciò è comprensibile dalle considerazioni più generali. Un'astrazione arbitraria, rimossa dalla sua base concreta (così come un numero è astratto da un insieme concreto di oggetti), non ha senso "in sé"; esiste solo nelle sue relazioni con altri concetti. Queste relazioni sono già implicite in ogni affermazione sull'astrazione, nella sua definizione più incompleta. Senza di esse l'astrazione manca di contenuto e significato, cioè semplicemente non esiste. Il contenuto del concetto di numero astratto sta nelle regole, nelle relazioni reciproche del sistema di numeri.

[6] Vale la pena notare che il concetto di numero, elaborato con tanta difficoltà in un lungo periodo di tempo, oggigiorno viene padroneggiato da un bambino con relativa facilità. Perché? La prima ragione, naturalmente, è che il bambino sente e vede gli adulti che usano costantemente i numeri e gli insegnano persino a fare lo stesso. Ma un secondo motivo, ed è quello su cui vogliamo richiamare particolare attenzione, è che il bambino ha già a portata di mano parole e simboli per i numeri. Prima impara questi simboli esterni per i numeri e solo più tardi ne padroneggia il significato.

[7] Il primo manoscritto indiano in cui compare lo zero risale alla fine del IX secolo; in esso il numero 270 è scritto esattamente come lo scriveremmo noi oggi. Ma è probabile che lo zero sia stato introdotto in India ancora prima, nel VI secolo.

[8] Ricordiamo che un numero primo è definito come un intero positivo maggiore dell'unità divisibile senza resto solo per il numero stesso e per l'unità.

[9] Qui si intendono i confini tra appezzamenti di terreno. Notiamo, tra parentesi, che geometria significa misura del territorio (in greco antico "ge" è terra, e "metron" è misura).

[10] Per forma si intendono anche le dimensioni.

[11] Confronta il capitolo XX.

[12] Questa teoria è attribuita allo scienziato greco Eudosso, vissuto nel IV secolo a.C.

[13] Poiché la teoria della misurazione delle grandezze non entrò a far parte dell'aritmetica ma passò alla geometria, presso i Greci la matematica fu inghiottita dalla geometria. Tali questioni, ad esempio, come la soluzione delle equazioni quadratiche, che oggi trattiamo in modo algebrico, sono state formulate e risolte geometricamente. Gli "Elementi" di Euclide contengono un numero considerevole di tali questioni, che ovviamente rappresentavano per i matematici contemporanei una sintesi dei fondamenti non solo della geometria nel nostro senso ma della matematica in generale. Questo dominio della geometria durò fino al tempo in cui Cartesio, al contrario, assoggettò la geometria all'algebra. Tracce del lungo dominio della geometria sono conservate, ad esempio, in nomi come "quadrato" e "cubo" per la seconda e la terza potenza: "a al cubo" è un cubo con il lato a.

[14] Non si tratta qui di una definizione descrittiva, ma di una definizione che serve come base immediata per dimostrazioni di teoremi sulle proprietà dei numeri reali. È naturale che tali definizioni siano nate in un periodo successivo, quando lo sviluppo della matematica, e in particolare dell'analisi infinitesimale, richiese un'adeguata definizione del numero reale rappresentato dalla "variabile x". Questa definizione fu data in varie forme negli anni settanta del secolo scorso dai matematici tedeschi Weierstrass, Dedekind e Cantor.

[15] Le frazioni con la cifra periodica nove non vengono qui considerate, sono identiche alla corrispondente frazione senza nove secondo la nota regola, che risulta chiara dall'esempio: 0,139999 ⋯ = 0,140000 ⋯.

[16] Tolomeo è ampiamente conosciuto come l'autore di un sistema in cui la Terra è considerata il centro dell'universo e il movimento dei corpi celesti è descritto come procedente attorno ad essa. Questo sistema venne soppiantato dal sistema copernicano.

[17] Fornisce le "equazioni" delle sezioni coniche riferite ad un vertice. Ad esempio "l'equazione" della parabola $y^2 = 2px$ è formulata così: Il quadrato di lato y ha un'area uguale al rettangolo di lati $2p$ e x. Naturalmente al posto dei simboli p, x, y vengono utilizzati i corrispondenti segmenti di linea.

[18] Per orientarci nelle date riportiamo le date di alcuni dei più importanti matematici dell'Oriente. Dall'India: Aryabhata, nato intorno al 476 d.C.; Brahmagupta, circa 598-660; Bhaskara, XII secolo; dalla Corasmia: Al-Kharizmi, IX secolo; Al-Biruni, 973-1048; dall'Azerbaigian: Nasireddin Tusi, 1201-74; da Samarcanda: Gyaseddin Jamschid, XV secolo.

[19] Bisogna tenere presente che è sbagliato associare lo sviluppo della matematica in questo periodo principalmente agli arabi. Il termine matematica "araba" venne utilizzato principalmente perché la maggior parte degli studiosi orientali scrivevano in lingua araba, che era stata diffusa in altri paesi dalle conquiste arabe.

[20] È da notare anche che il termine matematico "algoritmo", che denota un metodo o un insieme di regole per il calcolo, deriva dal nome dello stesso al-Kharizmi.

[21] È interessante notare che Napier non definì i logaritmi come vengono definiti oggi, quando diciamo che nella formula $x = a^y$ il numero y è il logaritmo di x in base a. Questa definizione di logaritmi è apparsa più tardi. La definizione di Napier era legata ai concetti di grandezza variabile e infinitesimo ed equivaleva a dire che il logaritmo di x è una funzione $y = f(x)$ il cui tasso di crescita è inversamente proporzionale a x; cioè $y' = c/x$ (vedi capitolo II). In questo modo la base della definizione era essenzialmente un'equazione differenziale, che definiva il logaritmo, sebbene i differenziali non fossero ancora stati inventati.

[22] La formula porta il nome di Newton non perché fu il primo a scoprirla ma perché la generalizzò da esponenti interi a esponenti frazionari arbitrari e irrazionali.

[23] I polinomi sono funzioni della forma $y = a_0x^n + a_1x^{n-1} + \ldots + a_n$. Il problema fondamentale dell'algebra del periodo, ovvero la soluzione dell'equazione $a_0x^n + a_1x^{n-1} + \ldots + a_n = 0$, significa semplicemente la ricerca dei valori di x per cui la funzione $y = a_0x^n + a_1x^{n-1} + \ldots + a_n$ è uguale a zero. L'esistenza stessa di una soluzione, di una radice dell'equazione, che viene chiamata teorema fondamentale dell'algebra, è dimostrata mediante l'analisi (vedi capitolo IV, §3).

[24] Questo duplice ruolo degli assiomi talvolta viene perso di vista anche in lavori di carattere metodologico, che attribuiscono così alla costruzione degli assiomi un significato che non le appartiene affatto, cioè quello della costruzione totale di una teoria.

[25] Nota dell'edizione inglese. Nel testo originale russo questa sezione è seguita da due sezioni intitolate «La natura essenziale della matematica» e «Le leggi dello sviluppo della matematica». Esse sono state omesse nella traduzione inglese perché discutono più dettagliatamente, e nel contesto filosofico più generale del materialismo dialettico, punti di vista già esposti con grande chiarezza nelle sezioni precedenti.

CAPITOLO II

Analisi

M. A. Lavrent'ev e S. M. Nikol'skij

§ 1. Introduzione

L'emergere, alla fine del Medioevo, di nuove condizioni manifatturiere in Europa, vale a dire la nascita del capitalismo, che in quel periodo stava sostituendo il sistema feudale, fu accompagnato da importanti scoperte ed esplorazioni geografiche. Nel 1492, basandosi sull'idea che la terra sia sferica, Colombo scoprì il Nuovo Mondo. La scoperta di Colombo allargò notevolmente i confini del mondo conosciuto e produsse una rivoluzione nella mente degli uomini. La fine del XV secolo e l'inizio del XVI videro l'attività creativa dei grandi artisti-umanisti Leonardo da Vinci, Raffaello e Michelangelo, che diedero un nuovo significato all'arte. Nel 1543 Copernico pubblicò la sua opera "Sulla rivoluzione dei corpi celesti", che cambiò completamente il volto dell'astronomia; nel 1609 apparve la "Nuova astronomia" di Keplero, contenente la sua prima e seconda legge sul moto dei pianeti attorno al sole, e nel 1618 il suo libro "Armonia del mondo", contenente la terza legge. Galileo, sulla base dello studio delle opere di Archimede e dei propri audaci esperimenti, gettò le basi della nuova meccanica, scienza indispensabile per la nascente tecnologia. Nel 1609 Galileo puntò verso il cielo notturno il suo cannocchiale di recente costruzione, benché ancora piccolo e imperfetto; bastò la prima occhiata al telescopio per distruggere le sfere celesti ideali di Aristotele e il dogma della forma perfetta dei corpi celesti. Si vide che la superficie della Luna era ricoperta di montagne e bucherellata di crateri. Venere mostrava fasi come la Luna, Giove era circondato da quattro satelliti e forniva un modello visivo in miniatura del sistema solare. La Via Lattea si divise in stelle separate e per la prima volta gli uomini percepirono l'incredibile distanza delle stelle. Nessun'altra scoperta scientifica ha mai lasciato una tale impressione sul mondo civilizzato.[1]

L'ulteriore sviluppo della navigazione, e di conseguenza dell'astronomia, nonché il nuovo sviluppo della tecnologia e della meccanica resero necessario lo studio di molti nuovi problemi matematici. La novità di questi problemi consisteva soprattutto nel fatto che richiedevano lo studio matematico delle leggi del moto nel senso lato del termine.

Lo stato di riposo e di immobilità è sconosciuto in natura. Tutta la natura, dalle particelle più piccole fino ai corpi più massicci, è in uno stato di eterna creazione e annientamento, in un flusso perpetuo, in incessante movimento e cambiamento. In ultima analisi, ogni scienza naturale studia qualche aspetto di questo movimento. L'analisi matematica è quella branca della matematica che fornisce metodi per l'indagine quantitativa di vari processi di cambiamento, movimento e dipendenza di una grandezza dall'altra. Sorse quindi naturalmente in un periodo in cui lo sviluppo della meccanica e dell'astronomia, animato dalle questioni della tecnologia e della navigazione, aveva già prodotto un notevole accumulo

di osservazioni, misurazioni e ipotesi e stava portando la scienza direttamente verso l'indagine quantitativa delle forme più semplici di movimento.

Il nome “analisi infinitesimale” non dice nulla sull’argomento in discussione ma enfatizza il metodo. Si tratta qui dello speciale metodo matematico degli infinitesimi o, nella sua forma moderna, dei limiti. Diamo ora alcuni esempi tipici di argomentazioni che fanno uso del metodo dei limiti e in uno dei paragrafi successivi definiremo i concetti necessari.

Esempio 1. Come stabilito sperimentalmente da Galileo, la distanza s percorsa nel tempo t da un corpo in caduta libera nel vuoto è espressa dalla formula

$$s = \frac{gt^2}{2} \qquad (1)$$

(g è una costante pari a 9,81 m/sec^2).[2] Qual è la velocità del corpo che cade in ogni punto del suo percorso?

Poniamo che il corpo passi per il punto A nell'istante t e consideriamo cosa accade nel breve intervallo di tempo di lunghezza Δt; cioè nel tempo da t a $t + \Delta t$. La distanza percorsa verrà aumentata di un certo incremento Δs. La distanza originale è $s_1 = gt^2/2$; la distanza aumentata è

$$s_2 = \frac{g(t+\Delta t)^2}{2} = \frac{gt^2}{2} + \frac{g}{2}(2t\Delta t + \Delta t^2).$$

Da questo troviamo l'incremento

$$\Delta s = s_2 - s_1 = \frac{g}{2}(2t\Delta t + \Delta t^2).$$

Questo rappresenta la distanza percorsa nel tempo da t a $t + \Delta t$. Per trovare la velocità media nel tratto di percorso Δs, dividiamo Δs per Δt:

$$v_{\text{av}} = \frac{\Delta s}{\Delta t} = gt + \frac{g}{2}\Delta t.$$

Lasciando che Δt si avvicini allo zero, otteniamo una velocità media che si avvicina quanto vogliamo alla velocità reale nel punto A. D'altra parte, vediamo che il secondo addendo a destra dell'equazione diventa estremamente piccolo al diminuire di Δt, per cui la media v_{av} si avvicina al valore gt, fatto che è conveniente scrivere come segue:

$$v = \lim_{\Delta t \to 0} v_{\text{av}} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \left(gt + \frac{g}{2}\Delta t\right) = gt.$$

Di conseguenza, gt è la velocità reale al momento t.

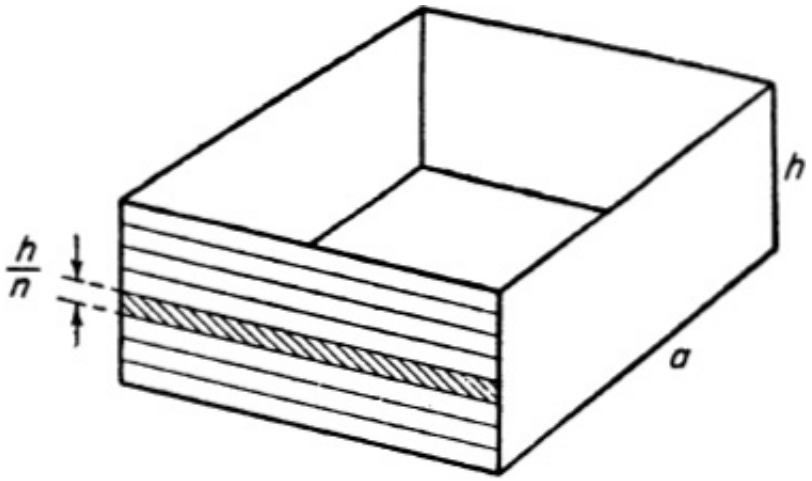


Figura 1.

Esempio 2. Un serbatoio con base quadrata di lato a e pareti verticali di altezza h è pieno d'acqua fino all'orlo (figura 1). Con quale forza agisce l'acqua su una delle pareti del serbatoio?

Dividiamo la superficie della parete in strisce orizzontali n di altezza h/n. La pressione esercitata in ciascun punto del recipiente è pari, secondo una legge ben nota, al peso della colonna d'acqua che vi giace sopra. Quindi sul bordo inferiore di ciascuna striscia la pressione, espressa in opportune unità, sarà rispettivamente pari a

$$\frac{h}{n}, \frac{2h}{n}, \frac{3h}{n}, \cdots, \frac{(n-1)\,h}{n}, h.$$

Otteniamo un'espressione approssimativa della forza desiderata P, se assumiamo che la pressione sia costante su ciascuna striscia. Pertanto il valore approssimativo di P è uguale a

$$P \approx \frac{ah}{n} \cdot \frac{h}{n} + \frac{ah}{n} \cdot \frac{2h}{n} + \cdots + \frac{ah}{n} \frac{(n-1)\,h}{n} + \frac{ah}{n} h$$

$$= \frac{ah^2}{n^2}(1 + 2 + \cdots + n) = \frac{ah^2}{n^2} \cdot \frac{n(n+1)}{2} = \frac{ah^2}{2}\left(1 + \frac{1}{n}\right).$$

Per trovare il vero valore della forza, dividiamo il lato in strisce sempre più strette, aumentando n senza limite. All'aumentare di n la grandezza $1/n$ nella formula sopra diventerà sempre più piccola e nel limite otteniamo la formula esatta

$$P = \frac{ah^2}{2}.$$

L'idea del metodo dei limiti è semplice e equivale a quanto segue. Per determinare il valore esatto di una certa grandezza, non determiniamo prima la grandezza stessa ma una certa approssimazione ad essa. Tuttavia non facciamo un'approssimazione ma tutta una serie di approssimazioni, ciascuna più precisa della precedente. Quindi dall'esame di questa catena di approssimazioni, cioè dall'esame del processo di approssimazione stesso, determiniamo in modo univoco il valore esatto della grandezza. Con questo metodo, che è essenzialmente

profondamente dialettico, otteniamo una costante fissa come risultato di un processo o movimento.

Il metodo matematico dei limiti si è evoluto come risultato del lavoro persistente di molte generazioni su problemi che non potevano essere risolti con i semplici metodi dell'aritmetica, dell'algebra e della geometria elementare.

Quali furono i problemi la cui soluzione portò ai concetti fondamentali dell'analisi, e quali furono i metodi di soluzione adottati per questi problemi? Esaminiamone alcuni.

I matematici del XVII secolo scoprirono gradualmente che un gran numero di problemi derivanti da vari tipi di movimento con conseguente dipendenza di alcune variabili da altre, e anche da problemi geometrici che non avevano ceduto ai metodi precedenti, potevano essere ridotti a due tipi. Semplici esempi di problemi del primo tipo sono: trovare la velocità in qualsiasi istante di un dato movimento non uniforme (o, più in generale, trovare la velocità di variazione di una data grandezza) e tracciare una tangente a una data curva. Questi problemi (il nostro primo esempio è uno di questi) portarono a un ramo dell'analisi che ricevette il nome di "calcolo differenziale". Gli esempi più semplici del secondo tipo di problemi sono: trovare l'area di una figura curvilinea (il problema della quadratura), oppure la distanza percorsa in un movimento non uniforme, o più in generale l'effetto totale dell'azione di grandezza continuamente variabile (confronta il secondo dei nostri due esempi). Questo gruppo di problemi portò ad un'altra branca dell'analisi, il "calcolo integrale". Furono così individuati due problemi fondamentali: il problema delle tangenti e il problema delle quadrature.

In questo capitolo descriveremo nel dettaglio le idee alla base della soluzione di questi due problemi. Particolarmente importante qui è il teorema di Newton e Leibniz secondo cui il problema delle quadrature è l'inverso, nel senso ben noto, del problema delle tangenti. Per risolvere il problema delle tangenti e i problemi ad esso riconducibili è stato elaborato un apposito algoritmo, un metodo del tutto generale che porta direttamente alla soluzione, cioè il metodo delle derivate o della derivazione.

La storia della nascita e dello sviluppo dell'analisi e del ruolo svolto nel suo sviluppo dalla geometria analitica di Cartesio è già stata descritta nel capitolo I. Vediamo che nella seconda metà del XVII secolo e nella prima metà del XVIII secolo si verificò un cambiamento completo in tutta la matematica. Alle divisioni già esistenti, l'aritmetica, la geometria elementare, i rudimenti dell'algebra e della trigonometria, si aggiunsero metodi generali come la geometria analitica, il calcolo differenziale e integrale e la teoria delle equazioni differenziali più semplici. Ora era possibile risolvere problemi le cui soluzioni fino ad allora erano state del tutto inaccessibili.

Si è scoperto che se la legge per la formazione di una determinata curva non è troppo complicata, allora è sempre possibile costruire una tangente ad essa in un punto arbitrario; è solo necessario calcolare, con l'aiuto delle regole del calcolo differenziale, la cosiddetta derivata, che nella maggior parte dei casi richiede tempi molto brevi. Fino ad allora era stato

possibile tracciare le tangenti solo al cerchio e ad una o due altre curve, e nessuno sospettava l'esistenza di una soluzione generale del problema.

Se conosciamo la distanza percorsa da un punto in movimento fino a un dato istante di tempo, con lo stesso metodo possiamo trovare immediatamente la velocità del punto in un dato momento, e anche la sua accelerazione. Dall'accelerazione è invece possibile ricavare la velocità e la distanza, sfruttando l'inverso della derivazione, cioè l'integrazione. Di conseguenza, non è stato molto difficile, ad esempio, dimostrare dalle leggi del movimento newtoniane e dalla legge della gravitazione universale che i pianeti devono muoversi attorno al sole in ellissi secondo le leggi di Keplero.

Della massima importanza nella vita pratica è il problema dei valori massimi e minimi di una grandezza, il cosiddetto problema dei massimi e dei minimi. Facciamo un esempio: Da un tronco di legno di sezione circolare di raggio dato vogliamo tagliare una trave di sezione rettangolare tale che offra la massima resistenza alla flessione. Quale dovrebbe essere il rapporto tra i lati? Un breve argomento sulla rigidezza delle travi di sezione trasversale rettangolare (applicando semplici concetti del calcolo integrale), seguito dalla risoluzione di un problema di massimo (che implica il calcolo di una derivata) fornisce la risposta che la rigidezza maggiore si produce per una sezione trasversale rettangolare la cui altezza è nel rapporto con la sua base di $\sqrt{2} : 1$. I problemi dei massimi e dei minimi si risolvono altrettanto semplicemente quanto quelli del disegno delle tangenti.

In vari punti di una linea curva, se non è una linea retta o un cerchio, la curvatura è in generale diversa. Come possiamo calcolare il raggio di un cerchio avente la stessa curvatura della linea data in un dato punto, il cosiddetto raggio di curvatura della curva nel punto? Si scopre che questo è altrettanto semplice; è necessario applicare l'operazione di derivazione solo due volte. Il raggio di curvatura gioca un ruolo importante in molte questioni di meccanica.

Prima dell'invenzione dei nuovi metodi di calcolo, era stato possibile trovare l'area solo dei poligoni, del cerchio, di un settore o segmento di cerchio, e di due o tre altre figure. Inoltre Archimede aveva già inventato un modo per calcolare l'area di un segmento di parabola. Il metodo estremamente ingegnoso da lui utilizzato in questo problema si basava su particolari proprietà della parabola e di conseguenza suscitò l'idea che ogni nuovo problema nel calcolo dell'area avrebbe molto probabilmente richiesto propri metodi di indagine, ancora più ingegnosi e difficili di quelli di Archimede. Perciò i matematici furono molto contenti quando si scoprì che il teorema di Newton e Leibniz, secondo cui l'inversione del problema delle tangenti avrebbe risolto il problema della quadratura, forniva subito un metodo per calcolare le aree delimitate da curve di tipo molto diverso. È diventato chiaro che esiste un metodo generale, adatto a un numero infinito di figure diverse. La stessa osservazione vale per il calcolo dei volumi, delle superfici, delle lunghezze delle curve, della massa dei corpi disomogenei, ecc.

Il nuovo metodo ha ottenuto risultati ancora maggiori nella meccanica. Sembrava che non ci fosse problema nella meccanica che i nuovi calcoli non avrebbero chiarito e risolto.

Non molto tempo prima Pascal aveva spiegato l'aumento delle dimensioni del vuoto di Torricelli all'aumentare dell'altitudine come conseguenza della diminuzione della pressione atmosferica. Ma qual è esattamente la legge che regola questa diminuzione? La risposta alla domanda è immediata studiando una semplice equazione differenziale.

È ben noto ai marinai che dovrebbero fare due o tre giri del cavo di ormeggio attorno all'argano se si vuole che un uomo sia in grado di tenere una grande nave all'ormeggio. Perché è questo? Si è scoperto che dal punto di vista matematico il problema è quasi del tutto identico al precedente e può essere risolto immediatamente.

Alla creazione dell'analisi seguì così un periodo di sviluppo tempestoso delle sue applicazioni ai più svariati rami della tecnologia e delle scienze naturali. Fondata sull'astrazione dalle particolarità dei problemi particolari, l'analisi matematica riflette le effettive proprietà profonde del mondo materiale; e questa è la ragione per cui fornisce i mezzi per indagare una gamma così ampia di questioni pratiche. Il movimento meccanico dei corpi solidi, il movimento dei liquidi e dei gas delle loro particolari particelle, le loro leggi di flusso nella massa, la conduzione del calore e dell'elettricità, il corso delle reazioni chimiche, tutti questi fenomeni vengono studiati nelle scienze corrispondenti mediante l'analisi matematica.

Nello stesso tempo in cui le sue applicazioni si estendevano, l'argomento stesso dell'analisi si arricchiva incommensurabilmente con la creazione e lo sviluppo di vari nuovi rami, come la teoria delle serie, le applicazioni della geometria all'analisi e la teoria delle equazioni differenziali.

Tra i matematici del XVIII secolo era diffusa l'opinione che qualsiasi problema delle scienze naturali, purché se ne trovasse una corretta descrizione matematica, potesse essere risolto mediante la geometria analitica e il calcolo differenziale e integrale.

I matematici procedettero gradualmente verso problemi più complicati di scienze naturali e tecnologia, che richiedevano un ulteriore sviluppo dei loro metodi. Per la soluzione di tali problemi divenne necessario creare ulteriori rami della matematica: il calcolo delle variazioni, la teoria delle funzioni di variabile complessa, la teoria dei campi, le equazioni integrali e l'analisi funzionale. Ma tutti questi nuovi metodi di calcolo erano essenzialmente estensioni e generalizzazioni immediate dei notevoli metodi scoperti nel XVII secolo. I più grandi matematici del XVIII secolo, David Bernoulli (1700-1782), Leonard Euler (1707-1783) e Lagrange (1736-1813), che aprirono nuove strade nella scienza, presero costantemente come punto di partenza i problemi fondamentali delle scienze esatte. Questo energico sviluppo dell'analisi fu continuato nel XIX secolo da matematici famosi come Gauss (1777-1855), Cauchy (1789-1857), M. V. Ostrogradskij (1801-1861), P. L. Čebyšev (1821-1894), Riemann (1826-1866), Abel (1802-1829), Weierstrass (1815-1897), i quali diedero tutti contributi davvero notevoli allo sviluppo dell'analisi matematica.

Il genio matematico russo N. I. Lobačevskiĭ ha avuto un influsso sullo sviluppo di alcune questioni dell'analisi matematica, ma vanno menzionati anche i principali matematici attivi all'inizio del XX secolo: A. A. Markov (1856-1922), A. M. Lyapunov (1857-1918), H. Poincaré (1854-1912), F. Klein (1849-1925), D. Hilbert (1862-1943).

La seconda metà del XIX secolo vide un profondo esame critico e una chiarificazione dei fondamenti dell'analisi. I vari metodi potenti che si erano accumulati furono ora posti su una base sistematica uniforme, corrispondente al livello avanzato di rigore matematico. Tutti questi metodi sono i mezzi con cui, insieme all'aritmetica, all'algebra, alla geometria e alla trigonometria, diamo un'interpretazione matematica del mondo che ci circonda, descriviamo il corso degli eventi reali e risolviamo gli importanti problemi pratici ad essi connessi.

La geometria analitica, il calcolo differenziale ed integrale, la teoria delle equazioni differenziali sono studiate in tutti gli istituti tecnici, tanto che queste branche della matematica sono conosciute da milioni di cittadini; gli elementi di queste scienze vengono insegnati anche in molte scuole tecniche; si discute anche della loro introduzione nelle scuole secondarie.

In tempi più recenti l'uso generalizzato di macchine calcolatrici rapide ha introdotto una nuova era nella matematica. Queste macchine, insieme ai rami della matematica appena menzionati, aprono nuove e strane possibilità per l'umanità.

Attualmente l'analisi e le branche che da essa derivano rappresentano una scienza matematica ampiamente diversificata, composta da numerose discipline indipendenti e strettamente connesse tra loro; ciascuna di queste discipline è in fase di sviluppo e perfezionamento.

Nell'analisi delle esigenze della vita quotidiana, dei problemi legati allo sviluppo imponente della tecnologia, giocano più che mai un ruolo significativo. Di grande importanza sono i problemi aerodinamici delle velocità ipersoniche, che vengono risolti con costante successo. I problemi più difficili della fisica matematica hanno ormai raggiunto lo stadio in cui possono essere risolti in forma numerica pratica. Nella fisica contemporanea teorie come la meccanica quantistica (che studia i problemi peculiari del microcosmo dell'atomo) non solo richiedono le branche più avanzate dell'analisi matematica contemporanea per risolvere i loro problemi, ma non potrebbero nemmeno descrivere i loro concetti fondamentali senza l'uso dell'analisi.

Lo scopo del presente capitolo è quello di fornire una presentazione divulgativa, adatta a un lettore che abbia familiarità solo con la matematica elementare, dello sviluppo e delle applicazioni più semplici di concetti basilari dell'analisi come funzione, limite, derivata e integrale. Poiché i vari rami particolari dell'analisi saranno trattati in altri capitoli del libro, il presente capitolo ha un carattere più elementare e un lettore che abbia già studiato un consueto primo corso di analisi può ometterlo senza pregiudicare la sua comprensione del resto del libro.

§ 2. Funzione

Il concetto di funzione. I vari oggetti o fenomeni che osserviamo in natura sono organicamente collegati tra loro; sono interdipendenti. Le relazioni più semplici di questo tipo sono note da tempo all'umanità e le informazioni su di esse sono state accumulate e formulate come leggi fisiche. Queste leggi indicano che le diverse grandezze che caratterizzano un dato fenomeno sono così strettamente legate tra loro che alcune di esse sono completamente determinate dai valori di altre. Ad esempio, la lunghezza dei lati di un rettangolo determina completamente la sua area, il volume di una data quantità di gas a una data temperatura è determinato dalla pressione e l'allungamento di una data barra metallica è determinato dalla sua temperatura. Sono state uniformità di questo tipo a dare origine al concetto di funzione.

Consideriamo una formula algebrica che, corrispondendo a ciascun valore delle grandezze letterali in essa occorrenti, ci permetta di trovare il valore della grandezza espressa dalla formula; l'idea di base qui è quella di una funzione. Consideriamo alcuni esempi di funzioni espresse da tali formule.

1. Supponiamo che all'inizio di un certo periodo di tempo un punto materiale fosse in riposo e che successivamente cominciasse a cadere per effetto della gravità. Allora la distanza s tracciata dal punto fino all'istante t è espressa dalla formula

$$s = \frac{gt^2}{2}, \qquad (1)$$

dove g è l'accelerazione di gravità.

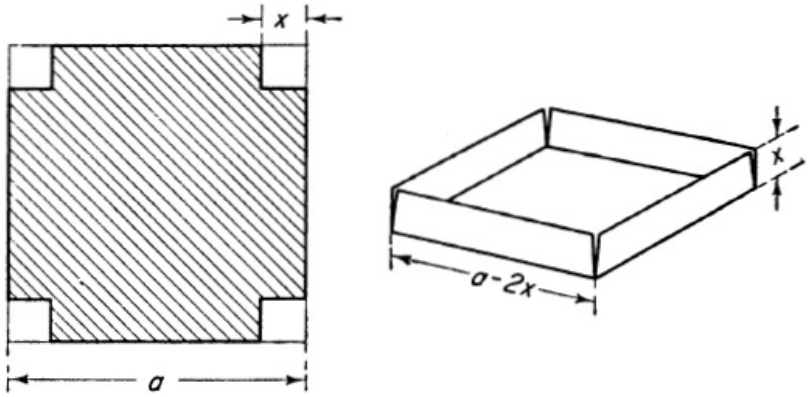


Figura 2.

2. Da un quadrato di lato a costruiamo una scatola rettangolare aperta di altezza x (figura 2). Il volume V della scatola viene calcolato dalla formula

$$V = x(a - 2x)^2. \qquad (2)$$

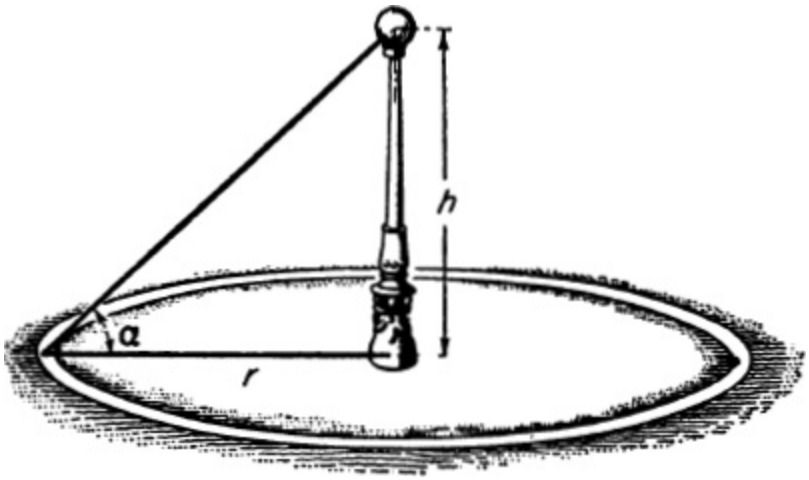


Figura 3.

La formula (2) ci permette, per ogni altezza x soggetta all'ovvio vincolo $0 \leqslant x \leqslant a/2$, di trovare il volume della scatola.

3. Si eriga un pilastro (figura 3) al centro di una pista di pattinaggio circolare con una luce all'altezza h. L'illuminazione T sul bordo del cerchio può essere espressa con la formula

$$T = \frac{A \sin \alpha}{h^2 + r^2}, \qquad (3)$$

dove r è il raggio del cerchio, $\tan \alpha = h/r$ e A è una certa grandezza che caratterizza la potenza della luce. Se conosciamo l'altezza h possiamo calcolare T dalla formula (3).

4. Le radici dell'equazione quadratica

$$x^2 + px - 1 = 0 \qquad (4)$$

sono dati dalla formula

$$x = -\frac{p}{2} \pm \sqrt{1 + \frac{p^2}{4}}. \qquad (5)$$

La caratteristica di una formula in generale, e degli esempi appena riportati in particolare, è che la formula permette, per ogni dato valore di una delle variabili (il tempo t, l'altezza x della scatola, l'altezza h del pilastro, il coefficiente p dell'equazione quadratica), detta variabile indipendente, di calcolare il valore dell'altra variabile (la distanza s, il volume V, l'illuminazione T, la radice x dell'equazione), che è chiamata variabile dipendente o funzione della prima variabile.

Ciascuna delle formule introdotte fornisce un esempio di funzione: la distanza s tracciata dal punto è funzione del tempo t; il volume V della scatola è funzione dell'altezza x; l'illuminazione T del bordo pista è funzione dell'altezza h del pilastro; le due radici dell'equazione quadratica (4) sono funzioni del coefficiente p.

È da notare che in alcuni casi la variabile indipendente può assumere qualsiasi valore numerico desiderato, come nell'esempio 4 dove il coefficiente p dell'equazione quadratica (4) può essere un numero arbitrario. In altri casi la variabile indipendente può assumere un valore arbitrario da qualche insieme (o raccolta) di numeri determinati in anticipo; come nell'esempio 2, dove il volume della scatola è funzione della sua altezza x, che può assumere qualsiasi valore dall'insieme dei numeri x che soddisfano la disuguaglianza $0 < x < a/2$. Allo stesso modo, nell'esempio 3 l'illuminazione T al bordo della pista è funzione dell'altezza h del pilastro, che teoricamente può assumere qualsiasi valore che soddisfi la disuguaglianza $h > 0$, ma in pratica h deve soddisfare le disuguaglianze $0 < h \leqslant H$, dove la grandezza H è determinato dalle strutture tecniche a disposizione dell'amministrazione della pista.

Introduciamo altri esempi di questo tipo. La formula

$$y = \sqrt{1 - x^2}$$

determina una funzione reale (che esprime una relazione tra i numeri reali x e y) solo per quei valori di x che soddisfano le disuguaglianze $-1 \leqslant x \leqslant +1$, e la formula $y = \log (1 - x^2)$ solo per quelli x che soddisfano le disuguaglianze $-1 < x < 1$.

È quindi necessario tenere conto del fatto che le funzioni effettive non possono essere definite per tutti i valori numerici della variabile indipendente ma solo per quei valori che appartengono a un determinato insieme, che molto spesso riempie un intervallo sull'asse x, con o senza punti finali.

Siamo ora in grado di dare la definizione di funzione accettata nella matematica odierna.

La grandezza (dipendente) y è funzione della grandezza (indipendente) x se esiste una regola per cui ad ogni valore di x appartenente ad un certo insieme di numeri corrisponde un determinato valore di y.

L'insieme di valori x che compaiono in questa definizione è chiamato dominio della funzione.

Ogni nuovo concetto dà origine a un nuovo simbolismo. Il passaggio dall'aritmetica all'algebra fu reso possibile dalla costruzione di formule valide per numeri arbitrari, e la ricerca di soluzioni generali diede origine al simbolismo letterale dell'algebra.

Il problema dell'analisi è lo studio delle funzioni, cioè della dipendenza di una variabile da un'altra. Di conseguenza, come nell'algebra si è verificato il passaggio dai numeri concreti ai numeri arbitrari, indicati con lettere, così nell'analisi si è verificato il corrispondente passaggio dalle formule concrete alle formule arbitrarie. La frase "y è una funzione di x" è convenzionalmente scritta come

$$y = f(x).$$

Come in algebra si usano lettere diverse per numeri diversi, così in analisi si usano notazioni diverse per tipi diversi di dipendenza, cioè per funzioni diverse: così scriviamo $y = F(x)$, $y = \phi(x)$, ⋯

Grafici di funzioni. Una delle idee più feconde e brillanti della seconda metà del XVII secolo fu l'idea del collegamento tra il concetto di funzione e la rappresentazione geometrica di una linea. Questo collegamento può essere realizzato, ad esempio, mediante un sistema di coordinate cartesiane rettangolare, con il quale il lettore ha sicuramente familiarità in modo generale dalla matematica della scuola secondaria.

Impostiamo sul piano un sistema di coordinate cartesiane rettangolare. Ciò significa che sul piano scegliamo due linee tra loro perpendicolari (l'asse delle ascisse e l'asse delle ordinate), su ciascuna delle quali fissiamo una direzione positiva. Allora a ciascun punto M del piano possiamo assegnare due numeri (x, y), che ne sono le coordinate, che esprimono nel sistema di misura dato la distanza, presa con il segno proprio, [3], del punto M rispettivamente dall'asse delle ordinate e dall'asse delle ascisse.

Con un tale sistema di coordinate possiamo rappresentare graficamente le funzioni sotto forma di determinate linee. Supponiamo che ci venga assegnata una funzione

$$y = f(x). \qquad (6)$$

Ciò significa, come sappiamo, che per ogni valore di x appartenente al dominio di definizione della funzione data, è possibile determinare in qualche modo, ad esempio mediante calcolo, un corrispondente valore y. Diamo a x tutti i possibili valori numerici, per ogni x determiniamo y secondo la nostra regola (6), e costruiamo sul piano il punto di coordinate x e y. In questo modo, ad ogni punto M' sull'asse x (figura 4) corrisponderà un punto M con coordinate x e $y = f(x)$. L'insieme di tutti i punti M forma una certa linea, che chiameremo grafico della funzione $y = f(x)$.

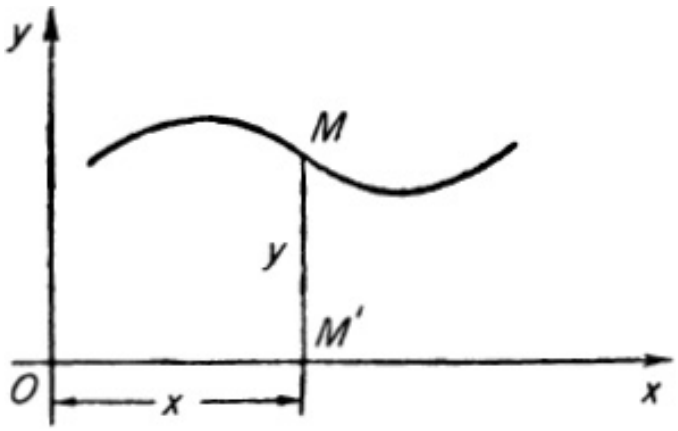


Figura 4.

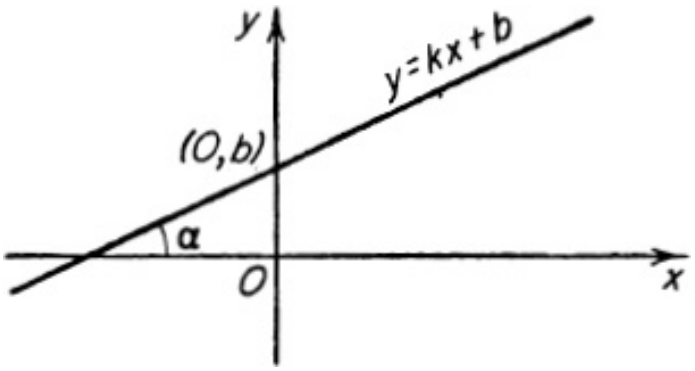


Figura 5.

Pertanto, il grafico della funzione f(x) è il luogo geometrico dei punti le cui coordinate soddisfano l'equazione (6).

A scuola abbiamo conosciuto i grafici delle funzioni più semplici. Quindi il lettore probabilmente sa che il grafico della funzione $y = kx + b$, dove k e b sono costanti, è una retta (figura 5) che forma l'angolo α con la direzione positiva dell'asse x, dove $\tan \alpha = k$ e interseca l'asse y nel punto (0, b). Questa funzione è chiamata funzione lineare.

Le funzioni lineari ricorrono molto frequentemente nelle applicazioni. Ricordiamo che molte leggi fisiche sono rappresentate, con notevole accuratezza, da funzioni lineari. Ad esempio, la lunghezza l di un corpo può essere considerata con buona approssimazione come funzione lineare della sua temperatura

$$l = l_0 + \alpha l_0 t,$$

dove α è il coefficiente di dilatazione lineare e l_0 è la lunghezza del corpo per $t = 0$. Se x è il tempo e y è la distanza percorsa da un punto in movimento, allora la funzione lineare $y = kx + b$ esprime ovviamente il fatto che il punto si muove con velocità uniforme k; e il numero b indica la distanza, al tempo $x_0 = 0$, del punto in movimento dal punto zero fisso da cui misuriamo le nostre distanze. Le funzioni lineari sono estremamente utili per la loro semplicità e perché è possibile considerare cambiamenti non uniformi come approssimativamente lineari, anche se solo per piccoli intervalli.

Ma in molti casi è necessario avvalersi della dipendenza funzionale non lineare. Ricordiamo ad esempio la legge di Boyle-Mariotte

$$v = \frac{c}{p},$$

dove le grandezze p e v sono inversamente proporzionali. Il grafico di tale relazione rappresenta un'iperbole (figura 6).

La legge fisica di Boyle-Mariotte corrisponde in realtà al caso in cui p e v sono positivi; rappresenta un ramo dell'iperbole che giace nel primo quadrante.

La classe generale dei processi oscillatori comprende i moti periodici, che di solito sono descritti dalle familiari funzioni trigonometriche. Ad esempio, se estendiamo una molla

sospesa dalla sua posizione di equilibrio, finché rimaniamo nei limiti elastici della molla, il punto *A* effettuerà delle oscillazioni verticali che sono espresse in modo abbastanza accurato dalla legge

$$x = a\cos(pt + \alpha),$$

dove *x* è lo spostamento del punto *A* dalla sua posizione di equilibrio, *t* è il tempo e i numeri *a*, *p* e *α* sono alcune costanti determinate dal materiale, dalle dimensioni e dall'estensione iniziale della molla.

Va tenuto presente che una funzione può essere definita in vari ambiti mediante varie formule, determinate dalle circostanze del caso. Ad esempio, la relazione $Q = f(t)$ tra la temperatura *t* di un grammo di acqua (o di ghiaccio) e la quantità di calore *Q* in essa contenuta, poiché *t* varia tra -10° e + 10°, è una funzione completamente determinata che è difficile esprimere in modo formula unica,[4] ma è facile rappresentare questa funzione con due formule. Poiché il calore specifico del ghiaccio è pari a 0,5 e quello dell'acqua è pari a 1, questa funzione, se concordiamo che *Q* = 0 a -10°. è rappresentato dalla formula

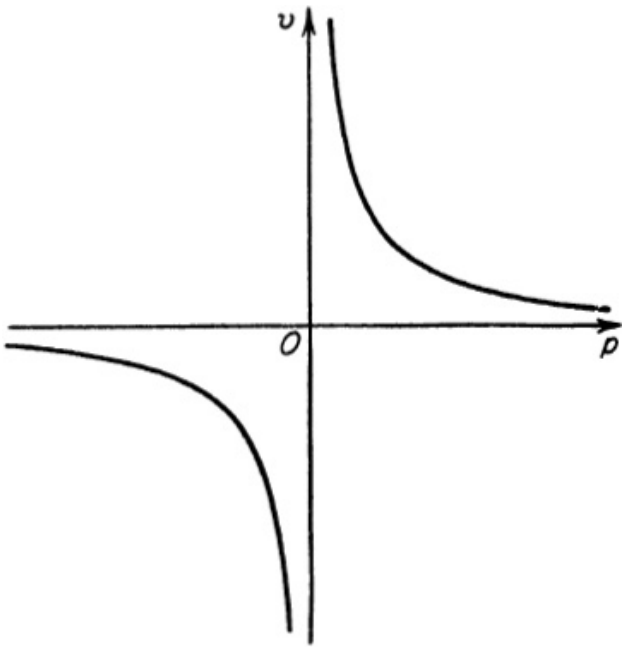


Figura 6

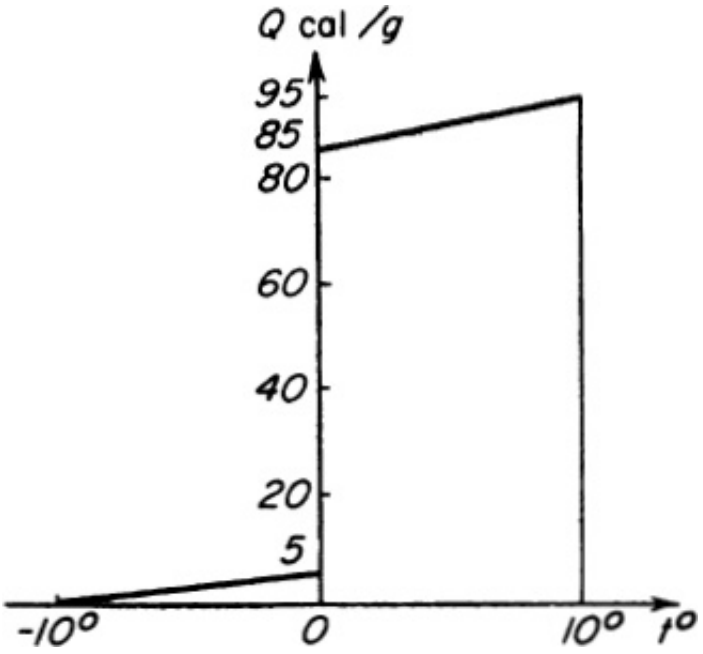


Figura 7

$$Q = 0.5t + 5,$$

poiché t varia nell'intervallo $-10^\circ \leqslant t < 0^\circ$ e con un'altra formula

$$Q = t + 85,$$

poiché t varia nell'intervallo $0^\circ < t \leqslant 10^\circ$. Per t = 0 questa funzione è indefinita o multivalore; per comodità, possiamo concordare che a t = 0 assume un valore ben definito, ad esempio $f(0) = 45$. Il grafico della funzione $Q = f(t)$ è riportato in figura 7.

Abbiamo introdotto molti esempi di funzioni date da formule. La possibilità di rappresentare una funzione mediante formule è estremamente importante dal punto di vista matematico, poiché tali formule forniscono condizioni molto favorevoli per studiare le proprietà delle funzioni con metodi matematici.

Ma non si deve pensare che una formula sia l'unico metodo per definire una funzione. Esistono molti altri metodi; ad esempio, il grafico della funzione, che ne fornisce un'immagine visiva geometrica. L'esempio seguente fornisce una buona illustrazione di un altro metodo.

Per registrare la variazione della temperatura dell'aria nel corso delle 24 ore, le stazioni meteorologiche si avvalgono di uno strumento chiamato termografo. Un termografo è costituito da un tamburo fatto ruotare attorno al proprio asse da un meccanismo a orologeria e da una struttura curva in ottone estremamente sensibile agli sbalzi di temperatura. Di conseguenza, una penna fissata al telaio tramite un sistema di leve si alza con l'aumentare della temperatura; e viceversa, un abbassamento della temperatura abbassa la penna. Sul tamburo è avvolto un nastro di carta millimetrata, sul quale la penna disegna una linea continua, formando il grafico della funzione $T = f(t)$, che esprime l'interdipendenza del tempo

e della temperatura dell'aria. Da questo grafico possiamo determinare, senza calcoli, il valore della temperatura in ogni istante *t*.

Questo esempio mostra che un grafico di per sé determina una funzione indipendentemente dal fatto che la funzione sia data da una formula o meno.

Per inciso, torneremo su questa questione (vedi Capitolo XII) e dimostreremo la seguente importante affermazione: ogni grafico continuo può essere rappresentato da una formula, o, come si usa ancora dire, da un'espressione analitica. Questa affermazione è vera anche per molti grafi discontinui.[5]

Notiamo che la verità di questa affermazione, di grande importanza teorica, è stata pienamente realizzata in matematica solo a metà del secolo scorso. Fino ad allora i matematici intendevano con il termine “funzione” solo un'espressione analitica (formula). Ma avevano l'errata impressione che molti grafici discontinui non corrispondessero ad alcuna espressione analitica, poiché presumevano che se una funzione era data da una formula, allora il suo grafico doveva possedere alcune proprietà particolarmente desiderabili rispetto agli altri grafici.

Ma nel XIX secolo si scoprì che ogni grafico continuo può essere rappresentato da una formula più o meno complicata. Così il ruolo eccezionale dell'espressione analitica come mezzo di definizione delle funzioni venne indebolito e venne alla luce la nuova definizione più flessibile data sopra per il concetto di funzione. Con questa definizione una variabile y si dice funzione di una variabile x se esiste una regola per cui ad ogni valore di x nel dominio di definizione della funzione corrisponde un valore y completamente determinato, indipendentemente dal modo in cui questa regola è data: mediante una formula, un grafico, una tabella o, in qualsiasi altro modo.

Possiamo qui osservare che nella letteratura matematica la suddetta definizione di funzione è spesso associata al nome di Dirichlet, ma vale la pena sottolineare che questa definizione è stata data simultaneamente e indipendentemente da N. I. Lobačevskiĭ. Infine suggeriamo come esercizio che il lettore disegni i grafici delle funzioni x^3, $\sqrt{x}$, $\sin x$, $\sin 2x$, $\sin(x + \pi/4)$, $\ln x$, $\ln(1 + x)$, $|x - 3|$, $(x + |x|)/2$.

Notiamo anche che il grafico di una funzione che per tutti i valori di x soddisfa la relazione

$$f(-x) = f(x)$$

è simmetrico rispetto all'asse y e nel caso

$$f(-x) = -f(x)$$

il grafico è simmetrico rispetto all'origine delle coordinate. Consideriamo anche come ottenere il grafico di una funzione $f(a + x)$, quando a è una costante, dal grafico di $f(x)$. Consideriamo infine come, utilizzando i grafici delle funzioni $f(x)$ e $\phi(x)$, sia possibile trovare i valori della funzione composta $y = f[\phi(x)]$.

§ 3. Limiti

Nel § 1 è stato affermato che l'analisi matematica moderna utilizza un metodo speciale, che è stato elaborato nel corso di molti secoli e che funge oggi da strumento fondamentale. Si tratta qui del metodo degli infinitesimi o, come è essenzialmente lo stesso, dei limiti. Cercheremo di dare qualche idea di questi concetti. A questo scopo consideriamo il seguente esempio.

Desideriamo calcolare l'area delimitata dalla parabola con l'equazione $y = x^2$, dall'asse x e dalla retta $x = 1$ (figura 8). La matematica elementare non ci fornirà i mezzi per risolvere questo problema. Ma ecco come possiamo procedere.

Dividiamo l'intervallo [0, 1] lungo l'asse x in parti uguali n nei punti

$$0, \frac{1}{n}, \frac{2}{n}, \cdots, \frac{n-1}{n}, 1$$

e su ciascuna di queste parti costruisci il rettangolo il cui lato sinistro si estende fino alla parabola. Si ottiene così il sistema di rettangoli ombreggiato in figura 8, la cui somma S_n delle cui aree è data da

$$S_n = 0 \cdot \frac{1}{n} + \left(\frac{1}{n}\right)^2 \frac{1}{n} + \left(\frac{2}{n}\right)^2 \frac{1}{n} + \cdots + \left(\frac{n-1}{n}\right)^2 \frac{1}{n}$$
$$= \frac{1^2 + 2^2 + \cdots + (n-1)^2}{n^3} = \frac{(n-1)\, n(2n-1)}{6n^3}.$$

Esprimiamo S_n nella forma seguente:

$$S_n = \frac{1}{3} + \left(\frac{1}{6n^2} - \frac{1}{2n}\right) = \frac{1}{3} + \alpha_n . \qquad (7)$$

La quantità α_n, che dipende da n, ha certamente un aspetto piuttosto ingombrante, ma possiede una certa proprietà notevole: se n viene aumentato oltre ogni limite, allora α_n si avvicina a 0. Questa proprietà può anche essere espressa come segue: se ci viene dato un numero positivo arbitrario ε, allora è possibile scegliere un numero intero N sufficientemente grande che per tutti gli n maggiori di N il numero α_n sarà inferiore al dato ε in valore assoluto.[7]

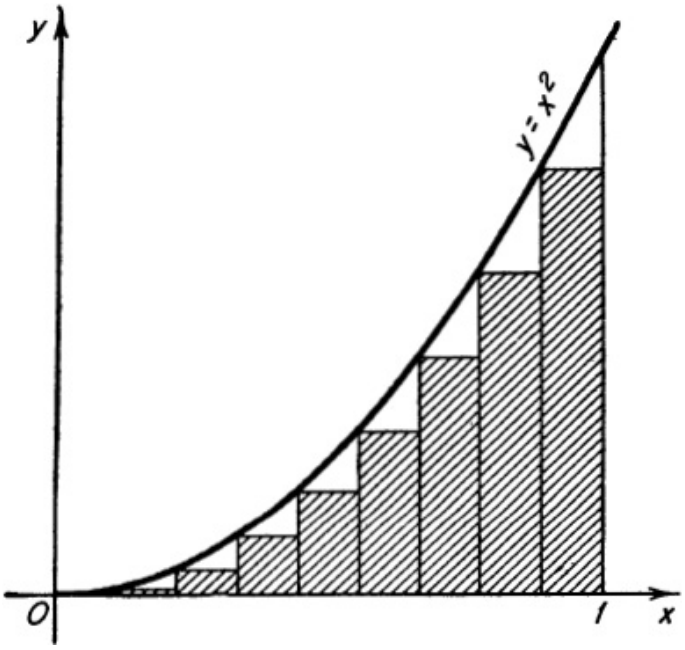


Figura 8.

La grandezza α_n è un esempio di infinitesimo nel senso in cui questa parola è usata nella matematica moderna.

Nella figura 8 vediamo che se aumentiamo il numero n oltre ogni limite, la somma S_n delle aree dei rettangoli ombreggiati si avvicinerà all'area desiderata della figura curvilinea. D'altra parte, l'equazione (7), in considerazione del fatto che α_n si avvicina a zero man mano che n aumenta oltre ogni limite, mostra che la somma S_n allo stesso tempo si avvicina a 1/3. Da ciò ne consegue che l'area desiderata S della figura è pari a 1/3, e abbiamo risolto il nostro problema.

Quindi il metodo in discussione equivale a questo, che per trovare una certa grandezza S introduciamo un'altra grandezza S_n, una grandezza variabile che si avvicina a S attraverso valori particolari $S_1, S_2, S_3, \cdots$, che dipendono secondo una legge dai numeri naturali $n = 1, 2, \cdots$. Quindi, dal fatto che la variabile S_n può essere rappresentata come la somma di una costante $\frac{1}{3}$ e di un infinitesimo α_n, concludiamo che S_n si avvicina a $\frac{1}{3}$ e quindi a $S = \frac{1}{3}$. Nel linguaggio della moderna teoria dei limiti possiamo dire che aumentando n la grandezza variabile S_n si avvicina ad un limite, che è uguale a $\frac{1}{3}$.

Diamo ora una definizione precisa dei concetti qui introdotti.

Se una grandezza variabile $\alpha_n (n = 1, 2, \cdots)$ ha la proprietà che per ogni numero positivo arbitrariamente piccolo ε è possibile scegliere un intero N così grande che per ogni $n > N$ abbiamo $|\alpha_n| < \varepsilon$, allora diciamo che α_n è un infinitesimo e scriviamo

$$\lim_{n\to\infty} \alpha_n = 0 \text{ or } \alpha_n \to 0.$$

D'altra parte, se una variabile x_n può essere rappresentata come una somma

$$x_n = a + \alpha_n ,$$

dove a è costante e α_n è un infinitesimo, allora diciamo che la variabile x_n, per n crescente oltre ogni limite, si avvicina al numero a e scriviamo

$$\lim x_n = a \text{ or } x_n \to a.$$

Il numero a è chiamato limite di x_n. In particolare il limite di un infinitesimo è ovviamente zero.

Consideriamo i seguenti esempi di grandezze variabili

$$x_n = \frac{1}{n} , y_n = -\frac{1}{n^2} , z_n = \frac{(-1)^n}{n} , u_n = \frac{n-1}{n} = 1 - \frac{1}{n} ;$$

$$v_n = (-1)^n \, (n = 1, 2, \cdots).$$

È chiaro che x_n, y_n e z_n sono infinitesimi, il primo si avvicina allo zero con valori decrescenti, il secondo con valori negativi crescenti, mentre il terzo assume valori che oscillano attorno allo zero. Inoltre $u_n \to 1$, mentre v_n non ha alcun limite, poiché al crescere di n non si avvicina ad alcun numero costante ma oscilla continuamente, assumendo i valori 1 e –1.

Un altro concetto importante in analisi è quello di grandezza infinitamente grande, che viene definita come variabile x_n (n = 1, 2, ⋯), con la proprietà che dopo aver scelto un numero positivo arbitrariamente grande M è possibile trovare un numero N tale che per tutti $n > N$

$$| x_n | > M.$$

Il fatto che la grandezza x_n è infinitamente grande si scrive così

$$\lim x_n = \infty \text{ or } x_n \to \infty.$$

Si dice che una grandezza x_n si avvicini all'infinito. Se è positivo (negativo) da qualche valore in poi, questo fatto è espresso così: $x_n \to +\infty$($x_n \to -\infty$). Ad esempio, per n = 1, 2, ⋯

$$\lim n^2 = +\infty, \lim (-n^3) = -\infty ;$$

$$\lim \log \frac{1}{n} = -\infty, \lim \tan \left(\frac{\pi}{2} + \frac{1}{n} \right) = -\infty.$$

È facile vedere che se una grandezza α_n è infinitamente grande, allora $\beta_n = 1/\alpha_n$ è infinitamente piccolo e viceversa.

Due grandezze variabili x_n e y_n possono essere sommate, sottratte, moltiplicate e divise l'una per l'altra in modo da produrre nuove grandezze che sono in generale anch'esse variabili: cioè la loro somma $x_n + y_n$, la loro differenza $x_n - y_n$, il loro prodotto $x_n y_n$ e il loro quoziente x_n/y_n. Di conseguenza saranno i loro valori particolari

$$x_1 \pm y_1 , x_2 \pm y_2 , x_3 \pm y_3 , \cdots$$

$$x_1 y_1 , x_2 y_2 , x_3 y_3 , \cdots$$

$$\frac{x_1}{y_1}, \frac{x_2}{y_2}, \frac{x_3}{y_3}, \cdots .$$

È anche possibile dimostrare, come è abbastanza evidente, che se le variabili x_n e y_n si avvicinano a limiti finiti, allora anche la loro somma, differenza, prodotto e quoziente si avvicina a limiti che sono corrispondentemente uguali alla somma, differenza, prodotto e quoziente di questi limiti. Questo fatto può essere espresso così:

$$\lim (x_n \pm y_n) = \lim x_n \pm \lim y_n ; \lim (x_n y_n) = \lim x_n \lim y_n ;$$

$$\lim \frac{x_n}{y_n} = \frac{\lim x_n}{\lim y_n} .$$

Tuttavia, nel caso del quoziente è necessario assumere che il limite del denominatore ($\lim y_n$) non sia uguale a zero. Se $\lim y_n = 0$ e $\lim x_n \neq 0$, il rapporto tra x_n e y_n non avrà un limite finito ma si avvicinerà all'infinito.

Particolarmente interessante e allo stesso tempo importante è il caso in cui il numeratore e il denominatore si avvicinano contemporaneamente allo zero. Qui non è possibile stabilire in anticipo se il rapporto x_n/y_n si avvicinerà a un limite e, in tal caso, quale sarà questo limite, poiché la risposta a questa domanda dipende interamente dal carattere dell'avvicinamento di x_n e y_n a zero. Ad esempio, se

$$x_n = \frac{1}{n} , y_n = \frac{1}{n^2} , z_n = \frac{(-1)^n}{n} (n = 1, 2, \cdots),$$

Poi

$$\frac{y_n}{x_n} = \frac{1}{n} \to 0, \frac{x_n}{y_n} = n \to \infty .$$

D'altra parte, la grandezza

$$\frac{x_n}{z_n} = (-1)^n$$

evidentemente non si avvicina ad alcun limite.

Quindi il caso in cui il numeratore e il denominatore della frazione si avvicinano entrambi allo zero non può essere trattato in anticipo con teoremi generali, e per ciascuna frazione particolare di questo tipo è necessario fare un'indagine speciale.

Vedremo più avanti che il problema fondamentale del calcolo differenziale, che può essere considerato come il problema di determinare la velocità di un movimento non uniforme in un

dato istante, si riduce a determinare il limite del rapporto tra due grandezze infinitesime, cioè l'aumento della distanza percorsa e l'aumento del tempo.

Finora abbiamo considerato le variabili x_n che assumono una successione di valori numerici $x_1, x_2, x_3, \cdots, x_n, \cdots$, mentre l'indice n percorre la successione dei numeri naturali $n = 1, 2, 3, \cdots$. Ma è anche possibile considerare il caso in cui n varia continuamente, come ad esempio il tempo, e qui anche determinare il limite della variabile x_n. Le proprietà di tali limiti sono del tutto analoghe a quelle formulate in precedenza per variabili discrete (cioè discontinue). Notiamo inoltre che non vi è alcun significato particolare nel fatto che n aumenti oltre ogni limite. È altrettanto possibile considerare il caso che, pur variando con continuità, n si avvicini ad un dato valore n_0.

Ad esempio, esaminiamo la variazione della grandezza di $(\sin x)/x$ quando x si avvicina allo zero. La tabella 1 mostra i valori di questa grandezza per alcuni valori di x:

Tabella 1

x	$\frac{\sin x}{x}$
0.50	0.9589...
0.10	0.9983...
0.05	0.9996...
...	...

(si presuppone che i valori di x siano espressi in radianti).

È ovvio che man mano che x si avvicina a zero la grandezza $(\sin x)/x$ si avvicina a 1, ma ovviamente dobbiamo ancora dare una dimostrazione rigorosa di questo fatto. La dimostrazione può essere ottenuta, ad esempio, dalla seguente disuguaglianza, che vale per tutti gli angoli diversi da zero nel primo quadrante:

$$\sin x < x < \tan x.$$

Se dividiamo entrambi i membri di questa disuguaglianza per il peccato x, otteniamo

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x},$$

da cui segue

$$\cos x < \frac{\sin x}{x} < 1.$$

Ma man mano che x diminuisce a zero il cos x si avvicina a 1, così che anche la grandezza $(\sin x)/x$, essendo contenuta nell'intervallo tra cos x e 1, si avvicina a 1, cioè

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

Avremo occasione più avanti di sfruttare questo fatto.

La nostra equazione è stata dimostrata nel caso in cui *x* si avvicina allo zero attraverso valori positivi. Ma modificando la dimostrazione in modo evidente, è possibile ottenere lo stesso risultato quando *x* si avvicina allo zero tramite valori negativi.

Discutiamo ora per un momento la seguente questione. Una grandezza variabile può avere o meno un limite e si pone la questione se sia possibile dare un criterio per determinare l'esistenza di un limite per una variabile. Ci limiteremo ad un caso importante e sufficientemente generale, per il quale si possa dare un simile criterio. Supponiamo che la grandezza variabile x_n aumenti o almeno non diminuisca; cioè, soddisfa le disuguaglianze

$$x_1 \leqslant x_2 \leqslant x_3 \leqslant \cdots,$$

e supponiamo anche di aver determinato che nessuno dei suoi valori supera un certo numero fisso *M*; cioè $x_n \leqslant M \ (n = 1, 2, \cdots)$. Se segniamo i valori di x_n e il numero *M* sull'asse *x*, vediamo che il punto variabile x_n si sposta lungo l'asse a destra ma rimane costantemente a sinistra del punto *M*. È piuttosto evidente che il punto variabile x_n deve inevitabilmente avvicinarsi ad un certo punto limite *a*, situato a sinistra di *M* o al massimo coincidente con *M*.

Quindi, nel caso in esame, il limite

$$\lim x_n = a$$

della nostra variabile esiste.

L'argomentazione di cui sopra ha carattere intuitivo ma possiamo considerarla come una dimostrazione. In un corso di analisi moderna viene data una dimostrazione completa di questo fatto sulla base della teoria dei numeri reali.

Consideriamo ad esempio la variabile

$$u_n = \left(1 + \frac{1}{n}\right)^n (n = 1, 2, 3, \cdots).$$

I primi valori sono $u_1 = 2$, $u_2 = 2{,}25$, $u_3 \approx 2{,}37$, $u_4 \approx 2{,}44$, ⋯, che si vedono aumentare. Dal teorema binomiale di Newton è possibile dimostrare che questo aumento vale per *n* arbitrario. Inoltre, è anche facile dimostrare che per tutti i *n* vale la disuguaglianza $u_n < 3$. Di conseguenza la nostra variabile deve avere un limite non maggiore di 3. Vedremo che questo limite gioca un ruolo molto importante nella fisica matematica e in un certo senso è la base più naturale per i logaritmi dei numeri.

È consuetudine indicare questo limite con la lettera *e*. È uguale a

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n = 2.718281828459045 \cdots.$$

Un'analisi più dettagliata mostra che il numero e non è razionale.[8]

È anche possibile dimostrare che il limite in esame esiste ed è pari a e non solo quando $n \to +\infty$ ma anche quando $n \to -\infty$. In entrambi i casi n può assumere anche valori non interi.

Citiamo un'importante applicazione alla fisica del concetto di limite. Consiste nel fatto notevole che solo utilizzando il concetto di limite (passaggio al limite) è possibile dare una definizione completa di molte delle grandezze concrete incontrate in fisica.

Consideriamo per il momento anche il seguente esempio geometrico. Nella geometria elementare le figure considerate per prime sono quelle delimitate da segmenti di retta. Ma poi si pone il compito più difficile di trovare la lunghezza della circonferenza di un cerchio di raggio dato.

Se analizziamo le difficoltà legate alla soluzione di questo problema, troviamo che si riducono a quanto segue.

Dobbiamo dare una risposta alla domanda, cosa si intende per lunghezza della circonferenza; dobbiamo cioè dare una definizione precisa di questa lunghezza. È essenziale che la definizione sia esprimibile in termini di lunghezze di segmenti rettilinei e che ci dia anche la possibilità di calcolare effettivamente la lunghezza della circonferenza.

Resta inteso, ovviamente, che il risultato di questo calcolo dovrebbe essere in accordo con l'esperienza pratica. Se ad esempio consideriamo una circonferenza costituita da un filo vero e proprio, allora, se tagliamo il filo e lo stiriamo, dobbiamo ottenere un segmento la cui lunghezza, nei limiti di accuratezza della misura, coincide con la lunghezza da noi calcolata.

Come noto dalla geometria elementare, la soluzione di questo problema si riduce alla seguente definizione. La lunghezza di una circonferenza è definita come il limite a cui si avvicina il perimetro di un poligono regolare[9] inscritto in esso quando il numero dei lati del poligono aumenta oltre tutti i limiti. La soluzione del problema si basa quindi essenzialmente sul concetto di limite.

La lunghezza di una curva uniforme arbitraria è definita allo stesso modo. Nei paragrafi che seguono incontreremo una serie di esempi di grandezze geometriche e fisiche che possono essere definite solo con il concetto di limite.

I concetti di limite e infinitesimo hanno trovato una formulazione definitiva all'inizio del secolo scorso. Le definizioni qui introdotte sono legate al nome di Cauchy, prima del quale i matematici operavano con concetti meno chiari. I concetti attuali di limite, di infinitesimo come grandezza variabile e di numero reale sono il risultato dello sviluppo dell'analisi matematica e sono stati allo stesso tempo il mezzo per affermare e chiarire le sue molteplici conquiste.

§ 4. Funzioni continue

Le funzioni continue costituiscono la classe di funzioni base per le operazioni di analisi matematica. L'idea generale di funzione continua si può ricavare dal fatto che il suo grafico è continuo; cioè la sua curva può essere disegnata senza staccare la matita dal foglio.

Una funzione continua dà l'espressione matematica di una situazione spesso riscontrata nella vita pratica, e cioè che ad un piccolo aumento di una variabile indipendente corrisponde un piccolo aumento della variabile dipendente, o funzione. Ottimi esempi di funzione continua sono dati dalle varie regole che governano il moto dei corpi $s = f(t)$, che esprimono la dipendenza della distanza s dal tempo t. Poiché il tempo e la distanza sono continui, una legge del moto del corpo $s = f(t)$ stabilisce tra loro una definita relazione continua, caratterizzata dal fatto che ad un piccolo aumento del tempo corrisponde un piccolo aumento della distanza.

L'umanità è arrivata all'astrazione della continuità osservando i cosiddetti mezzi densi circostanti, vale a dire solidi, liquidi e gas; ad esempio metalli, acqua e aria. In realtà, come è ormai noto, ogni mezzo fisico rappresenta l'accumulo di un gran numero di particelle separate in movimento. Ma queste particelle e le distanze tra loro sono così piccole in confronto alle dimensioni del mezzo in cui hanno luogo i fenomeni della fisica microscopica, che molti di questi fenomeni possono essere studiati con sufficiente accuratezza se consideriamo il mezzo approssimativamente privo di interstizi, cioè continuamente distribuito nello spazio occupato. È su tale presupposto che si basano molte scienze fisiche, ad esempio l'idrodinamica, l'aerodinamica e la teoria dell'elasticità. Il concetto matematico di continuità gioca naturalmente un ruolo importante in queste scienze, e anche in molte altre.

Consideriamo una funzione arbitraria $y = f(x)$ e un valore specifico della variabile indipendente x_0. Se la nostra funzione riflette un processo continuo, allora a valori x che differiscono solo leggermente da x_0 corrisponderanno valori della funzione $f(x)$ che differiscono solo leggermente dal valore $f(x_0)$ nel punto x_0. Pertanto, se l'incremento $x - x$ della variabile indipendente è piccolo, anche l'incremento corrispondente $f(x) - f(x_0)$ della funzione sarà piccolo. In altre parole se l'incremento della variabile indipendente $x - x_0$ si avvicina a zero, allora anche l'incremento $f(x) - f(x_0)$ della funzione deve avvicinarsi a zero, fatto che può essere espresso in nel modo seguente:

$$\lim_{x-x_0\to 0} [f(x) - f(x_0)] = 0. \qquad (8)$$

Questa relazione costituisce la definizione matematica di continuità della funzione nel punto x_0; vale a dire, la funzione $f(x)$ si dice continua nel punto x_0, se vale l'uguaglianza (8).

Infine diamo la seguente definizione. Una funzione si dice continua in un dato intervallo se è continua in ogni punto x_0 di tale intervallo; cioè, se in ciascuno di questi punti l'uguaglianza (8) è soddisfatta.

Pertanto, per introdurre una definizione matematica della proprietà di una funzione riflessa nel fatto che il suo grafico è continuo (nel senso quotidiano di questa parola), è stato

necessario prima definire la continuità locale (continuità nel punto x_0) e poi su questa base definire la continuità della funzione nell'intero intervallo.

Questa definizione, introdotta per la prima volta all'inizio del secolo scorso da Cauchy, è oggi generalmente adottata nell'analisi matematica contemporanea. Il test di molti esempi concreti ha dimostrato che corrisponde molto bene alla nozione pratica che abbiamo di una funzione continua, ad esempio, come rappresentata dal suo grafico continuo.

Come esempi di funzioni continue, il lettore può considerare le funzioni elementari a lui ben note dalla matematica scolastica x^n, sin x, cos x, come, log x, arc sin x, arc cos x. Tutte queste funzioni sono continue negli intervalli per i quali sono definite.

Se si sommano, sottraggono, moltiplicano o dividono funzioni continue (ad eccezione della divisione per zero), anche il risultato è una funzione continua. Ma nel caso della divisione la continuità viene solitamente distrutta per quei valori x_0 per i quali la funzione al denominatore si annulla. Il risultato della divisione in questo caso è una funzione discontinua nel punto x_0.

La funzione $y = 1/x$ può servire come esempio di funzione discontinua nel punto $x = 0$. Altre funzioni discontinue sono rappresentate dai grafici di figura 9.

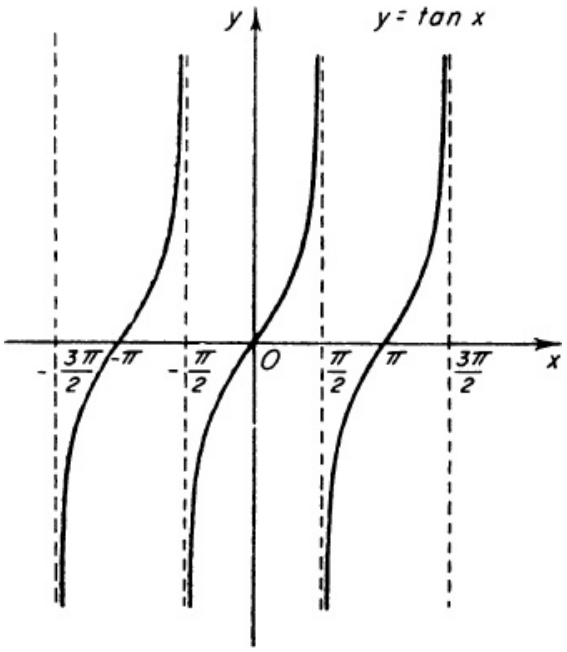


Figura 9a.

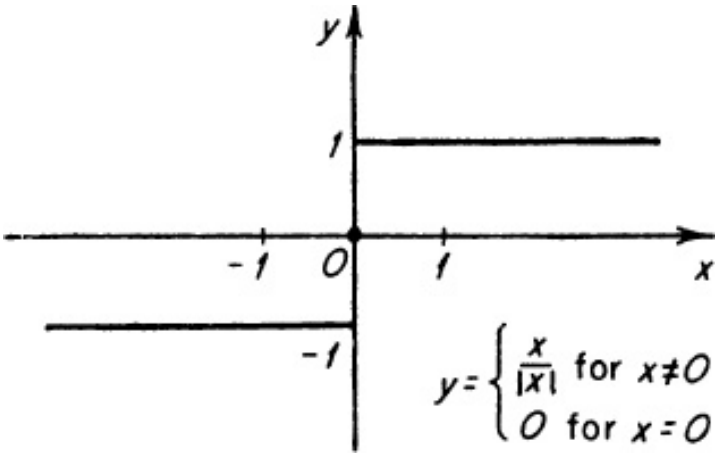


Figura 9b. («for»: per).

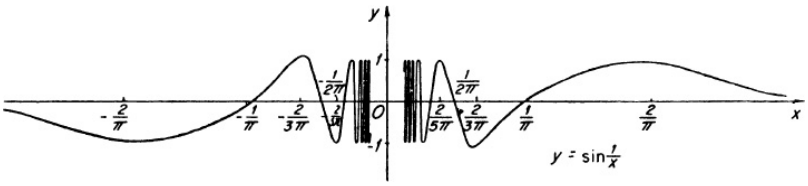


Figura 9c.

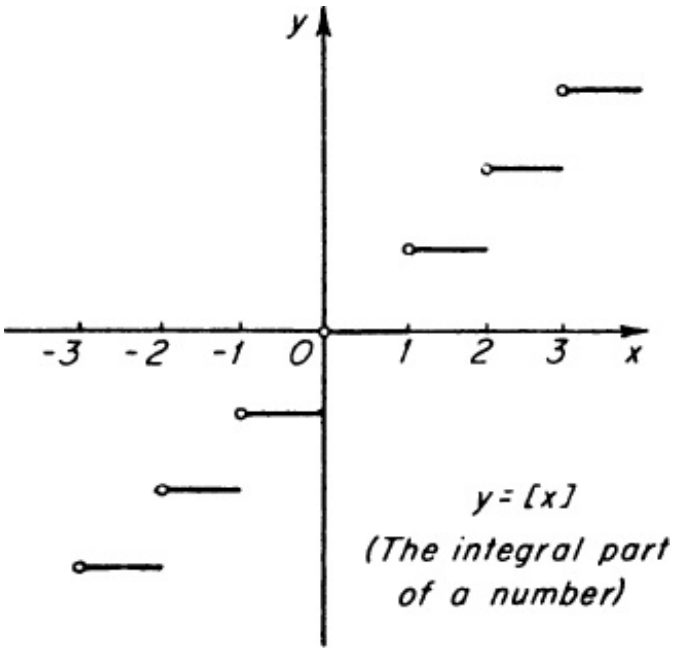


Figura 9d. Parte intera di un numero.

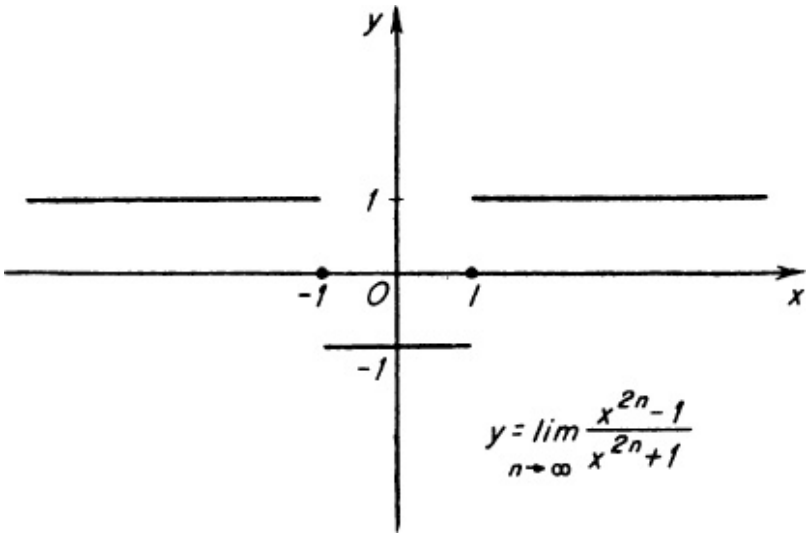


Figura 9e.

Raccomandiamo al lettore di esaminare attentamente questi grafici. Noterà che le interruzioni nelle funzioni sono di diverso tipo: In alcuni casi esiste un limite $f(x)$ quando x si avvicina al punto x_0 dove la funzione subisce una discontinuità, ma questo limite è diverso da $f(x_0)$. In altri casi, come nella figura 9c, il limite semplicemente non esiste. Può anche succedere che quando x si avvicina a x_0 da un lato $f(x) - f(x_0) \to 0$, ma quando $x \to x_0$ dall'altro lato, $f(x) - f(x_0)$ non si avvicina allo zero. In questo caso, ovviamente, la funzione presenta una discontinuità, ma possiamo dire che in quel punto è "continua da un lato". Tutti questi casi sono rappresentati nei grafici di figura 9.

Come esercizio raccomandiamo al lettore di porsi la domanda: quale valore bisogna dare alle funzioni

$$\frac{\sin x}{x}, \frac{1-\cos x}{x^2}, \frac{x^3-1}{x-1}, \frac{\tan x}{x}$$

nei punti in cui non sono definiti (cioè nei punti in cui il denominatore è uguale a zero), affinché possano essere continui in questi punti. Inoltre, è possibile trovare tali numeri per le funzioni

$$\tan x, \frac{1}{x-1}, \frac{x-2}{(x^2-4)} ?$$

Queste funzioni discontinue in matematica rappresentano i numerosi processi simili a salti che si incontrano in natura. Nel caso di un colpo improvviso, ad esempio, il valore della velocità di un corpo cambia in modo simile a un salto. Molte transizioni qualitative avvengono con tali salti. Nel §2 abbiamo introdotto la funzione $Q = f(t)$, che esprime il modo in cui la quantità di calore in una data quantità di acqua (o ghiaccio) dipende dalla temperatura. Nelle vicinanze del punto di fusione del ghiaccio la quantità di calore $Q = f(t)$ cambia in modo saltellante al variare di t.

Funzioni con discontinuità isolate si incontrano abbastanza spesso nell'analisi, insieme alle funzioni continue. Ma come esempio di funzione più complicata, dove il numero di discontinuità è infinito, consideriamo la cosiddetta funzione di Riemann, che è uguale a zero in tutti i punti irrazionali e uguale a $1/q$ nei punti razionali della forma $x = p/q$ (dove p/q è una frazione ai suoi termini più bassi). Questa funzione è discontinua in tutti i punti razionali e continua nei punti irrazionali. Modificandolo leggermente possiamo facilmente ottenere un esempio di una funzione discontinua in tutti i suoi punti. [10] Notiamo tra l'altro che anche per funzioni così complicate l'analisi moderna ha scoperto molte leggi interessanti, che vengono studiate in una delle branche indipendenti dell'analisi, la teoria delle funzioni di variabile reale. Questa teoria si è sviluppata con straordinaria rapidità negli ultimi 50 anni.

§ 5. Derivata

Il successivo concetto fondamentale dell'analisi è il concetto di derivata. Consideriamo due problemi da cui storicamente è nata.

Velocità. All'inizio di questo capitolo abbiamo definito la velocità di un corpo in caduta libera. Per fare ciò ci siamo avvalsi di un passaggio al limite dalla velocità media su brevi distanze alla velocità nel punto dato e nel tempo dato. La stessa procedura può essere utilizzata per definire la velocità istantanea per un movimento arbitrario non uniforme. In effetti, lasciamo la funzione

$$s = f(t) \qquad (9)$$

esprimere la dipendenza della distanza s percorsa dal punto materiale nel tempo t. Per trovare la velocità nell'istante $t = t_0$, consideriamo l'intervallo di tempo da t_0 a $t_0 + h (h \neq 0)$. Durante questo tempo il punto coprirà la distanza

$$\Delta s = f(t_0 + h) - f(t_0) .$$

La velocità media v_{av} su questa parte del percorso dipenderà da h

$$v_{av} = \frac{\Delta s}{h} = \frac{1}{h}[f(t_0 + h) - f(t_0)],$$

e rappresenterà la velocità effettiva nel punto t_0 con una precisione sempre maggiore man mano che h diventa più piccolo. Ne consegue che la velocità reale nel momento t_0 è pari al limite

$$v = \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h}$$

del rapporto tra l'aumento della distanza e l'aumento del tempo, poiché quest'ultimo si avvicina allo zero senza mai essere effettivamente uguale a zero. Per calcolare la velocità per diverse forme di movimento, dobbiamo scoprire come trovare questo limite per varie funzioni $f(t)$.

Tangente. Un altro problema, questa volta geometrico, ci porta a indagare un limite esattamente analogo, vale a dire il problema di tracciare una tangente a una curva piana arbitraria.

Sia la curva C il grafico di una funzione $y = f(x)$, e sia A il punto sulla curva C con ascissa x_0 (figura 10). Quale retta chiameremo tangente a C nel punto A? Nella geometria elementare questa questione non si pone. L'unica curva lì studiata, cioè la circonferenza di un cerchio, permette di definire la tangente come una retta che ha un solo punto in comune con la curva. Ma per altre curve tale definizione chiaramente non corrisponderà alla nostra immagine intuitiva di "tangenza". Pertanto, delle due rette L e M della figura 11, la prima ovviamente non è tangente alla curva ivi tracciata (una curva sinusoidale), sebbene abbia con essa un solo

punto in comune; mentre la seconda retta ha molti punti in comune con la curva, eppure è tangente alla curva in ciascuno di questi punti.

Per definire la tangente consideriamo sulla curva *C* (figura 10) un altro punto *A′*, distinto da *A*, con ascissa $x_0 + h$. Disegniamo la secante AA′ e indichiamo con α l'angolo che forma con l'asse *x*. Lasciamo ora che il punto *A′* si avvicini a *A* lungo la curva *C*. Se la secante AA′ si avvicina corrispondentemente ad una posizione limite, allora la retta *T* che ha questa posizione limite si chiama tangente nel punto *A*. Evidentemente l'angolo α formato dalla retta *T* con l'asse *x*, deve essere uguale al valore limite dell'angolo variabile β.

Il valore di tan β è facilmente determinabile dal triangolo ABA′ (figura 10):

$$\tan\beta = \frac{BA'}{AB} = \frac{f(x_0 + h) - f(x_0)}{h}.$$

Per la posizione limitante che dobbiamo avere

$$\tan\alpha = \lim_{A' \to A} \tan\beta = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

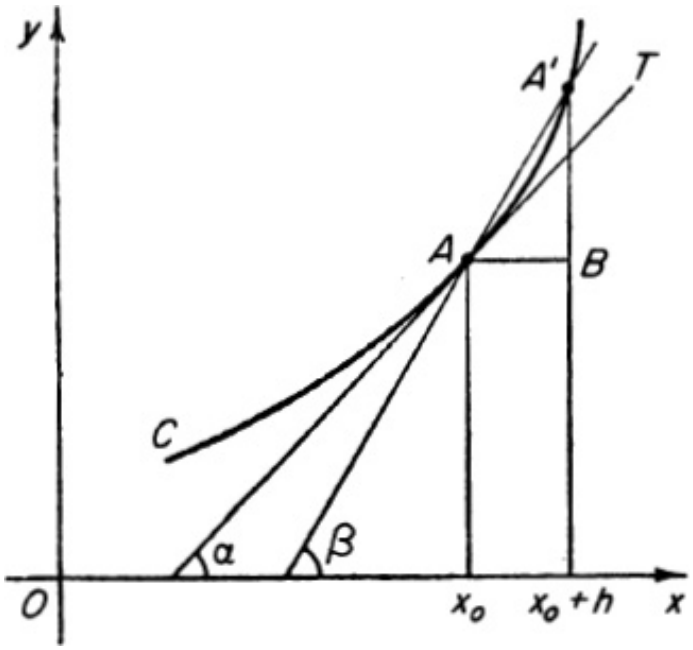


Figura 10.

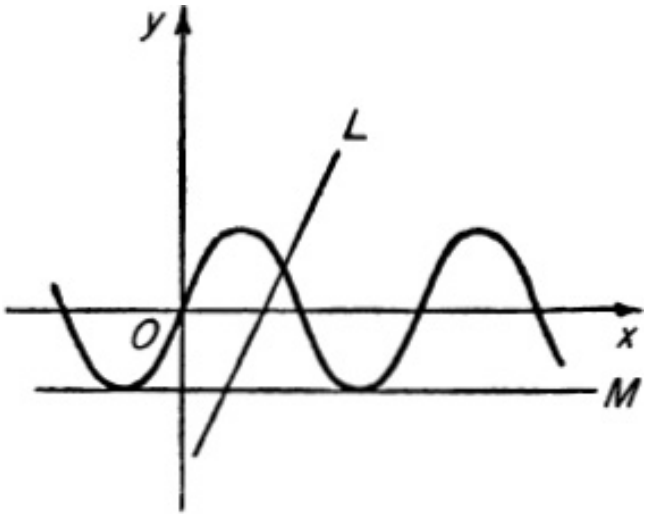


Figura 11.

cioè la tangente trigonometrica dell'angolo di inclinazione della linea tangente è uguale al limite del rapporto tra l'incremento della funzione $f(x)$ nel punto x_0 e il corrispondente incremento della variabile indipendente, poiché quest'ultima si avvicina a zero senza mai essere effettivamente uguale a zero.

Diamo ancora un altro esempio che porti al calcolo di un limite analogo. Supponiamo che una corrente elettrica variabile scorra attraverso un conduttore. Supponiamo di conoscere la funzione $Q = f(t)$ che esprime la quantità di energia elettrica che è passata attraverso una sezione fissa del conduttore fino all'istante t. Nel periodo da t_0 a $t_0 + h$, attraverso questa sezione scorrerà una quantità di energia elettrica ΔQ pari a $f(t_0 + h) - f(t_0)$. Il valore medio della corrente sarà quindi pari a

$$I_{av} = \frac{\Delta Q}{h} = \frac{f(t_0 + h) - f(t_0)}{h}.$$

Il limite di questo rapporto come $h \to 0$ ci darà il valore della corrente al momento t_0

$$I = \lim_{h \to 0} \frac{f(t_0 + h) - f(t_0)}{h}.$$

Tutti e tre i problemi discussi, nonostante si riferiscano a rami diversi della scienza, vale a dire la meccanica, la geometria e la teoria dell'elettricità, hanno portato a una stessa operazione matematica da eseguire su una data funzione, cioè a trovare il limite del rapporto tra l'aumento della funzione e il corrispondente aumento h della variabile indipendente come $h \to 0$. Il numero di problemi così diversi potrebbe essere aumentato a piacimento e la loro soluzione porterebbe alla stessa operazione. A essa siamo condotti, ad esempio, dalla questione della velocità di una reazione chimica, o della densità di una massa non omogenea e così via. In considerazione del ruolo eccezionale svolto da questa operazione sulle funzioni, ha ricevuto un nome speciale, derivazione, e il risultato dell'operazione è chiamato derivata della funzione.

La derivata della funzione $y = f(x)$, più precisamente il suo valore nel punto x, è dunque il limite[11] del rapporto fra l'incremento $f(x + h) - f(x)$ della funzione e l'incremento h della variabile indipendente, quando h tende a zero. Spesso si scrive $h = \Delta x$ e $f(x + \Delta x) - f(x) = \Delta y$; la definizione della derivata assume allora la forma concisa:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Il valore della derivata dipende ovviamente dal punto x in cui si trova. Pertanto la derivata di una funzione $y = f(x)$ è essa stessa una funzione di x. È consuetudine denotare la derivata in questo modo

$$f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

Anche alcune altre notazioni sono consuete per la derivata:

$$\frac{df(x)}{dx}, \text{ or } \frac{dy}{dx}, \text{ or } y', \text{ or } y'_x.$$

Da notare inoltre che la notazione $\frac{dy}{dx}$ appare come una frazione, sebbene venga letta come un unico simbolo della derivata. Nei paragrafi successivi il numeratore e il denominatore di questa "frazione" assumeranno significato autonomo, in modo tale che il loro rapporto coinciderà con la derivata così che questo modo di scrivere è pienamente giustificato.

I risultati di questi esempi possono ora essere formulati come segue.

La velocità di un punto per il quale la distanza s è una funzione data del tempo $s = f(t)$ è uguale alla derivata di questa funzione

$$v = s' = f'(t).$$

Più brevemente, la velocità è la derivata della distanza rispetto al tempo.

La tangente trigonometrica dell'angolo di inclinazione della retta tangente alla curva $y = f(x)$ nel punto con ascissa x è uguale alla derivata della funzione $f(x)$ in questo punto:

$$\tan \alpha = y' = f'(x).$$

L'intensità della corrente I al momento t, se $Q = f(t)$ è la quantità di elettricità che fino al momento t è passata attraverso una sezione trasversale del conduttore, è pari alla derivata

$$I = Q' = f'(t).$$

Facciamo la seguente osservazione. La velocità di un movimento non uniforme in un dato istante è un concetto puramente fisico, derivante dall'esperienza pratica. L'umanità vi è arrivata in seguito a numerose osservazioni su diversi movimenti concreti. Lo studio del movimento non uniforme di un corpo in diversi punti del suo percorso, il confronto di diversi

movimenti di questo tipo che si verificano simultaneamente, e in particolare lo studio dei fenomeni di collisione dei corpi, hanno rappresentato un accumulo di esperienza pratica che ha portato alla formulazione del concetto fisico di velocità di un movimento non uniforme in un dato momento. Ma la definizione esatta della velocità dipendeva necessariamente dal metodo con cui si definiva il suo valore numerico, e definire questo valore era possibile solo con il concetto di derivata.

In meccanica la velocità di un corpo che si muove secondo la regola $s = f(t)$ all'istante t è definita come la derivata della funzione $f(t)$ per questo valore di t.

La discussione all'inizio di questa sezione ha mostrato, da un lato, i vantaggi dell'introduzione dell'operazione di ricerca della derivata, e dall'altro ha fornito una ragionevole giustificazione per la definizione di velocità in un dato momento sopra formulata.

Pertanto, quando ci siamo posti il problema di trovare la velocità di un punto in movimento non uniforme, avevamo, propriamente parlando, solo una nozione empirica del suo valore, ma nessuna definizione esatta. Ma ora, come risultato della nostra analisi, siamo arrivati ad una definizione esatta del valore della velocità in un dato istante, cioè la derivata della distanza rispetto al tempo. Questo risultato è estremamente importante dal punto di vista pratico, poiché la nostra conoscenza empirica della velocità è stata notevolmente arricchita dal fatto che ora possiamo effettuare un calcolo numerico esatto.

Ciò che è stato appena detto si riferisce altrettanto bene, naturalmente, alla forza di una corrente e a molti altri concetti che esprimono la velocità di qualche processo, fisico, chimico, ecc.

Questa situazione può servire da esempio per numerose altre di natura simile, dove l'esperienza pratica ha portato alla formazione di un concetto relativo al mondo esterno (velocità, lavoro, densità, superficie, ecc.) e poi la matematica ci ha permesso di definire con precisione questo concetto, dopo di che possiamo utilizzare il concetto nei calcoli pratici.

Abbiamo già notato all'inizio del capitolo che il concetto di derivata è nato principalmente come risultato di molti secoli di sforzi diretti alla soluzione di due problemi: disegnare una tangente a una curva e trovare la velocità di un movimento non uniforme. Questi problemi, ed anche il calcolo delle aree discusso più avanti, interessarono i matematici dell'antichità. Ma fino al XVI secolo l'enunciato e il metodo di soluzione di ciascun problema di questo genere avevano un carattere del tutto particolare. L'accumulo di tutto questo vasto materiale fu ridotto a un sistema teoricamente completo nel XVII secolo grazie ai lavori di Newton e Leibniz. Un importante contributo ai fondamenti dell'analisi odierna è stato dato anche da Eulero.

Ma va detto che Newton e Leibniz e i loro contemporanei fornirono ben poche basi logiche per le loro grandi scoperte matematiche; nei loro metodi di ragionamento e nei concetti con cui operavano c'era molto di poco chiaro dal nostro punto di vista. Anche allora gli stessi matematici ne erano ben consapevoli, come dimostrano le aspre discussioni che si trovano nella loro corrispondenza. Tuttavia questi matematici dei secoli XVII e XVIII esercitarono la loro attività puramente matematica in strettissimo legame con le ricerche di altri ricercatori,

nei vari rami delle scienze naturali (fisica, meccanica, chimica, tecnologia). L'enunciazione di un problema matematico nasceva solitamente da esigenze pratiche o dal desiderio di comprendere qualche fenomeno della natura, e non appena il problema veniva risolto, la soluzione veniva sottoposta in un modo o nell'altro a una prova pratica. Di conseguenza, nonostante una certa mancanza di basi logiche, la matematica ha potuto avanzare in direzioni estremamente utili.

Esempi per il calcolo delle derivate. La definizione di derivata come limite

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

ci permette di calcolare la derivata di una determinata funzione concreta.

Naturalmente bisogna ammettere che sono possibili casi in cui la funzione in un punto o nell'altro o anche in molti punti semplicemente non ha derivata; in altre parole, il rapporto

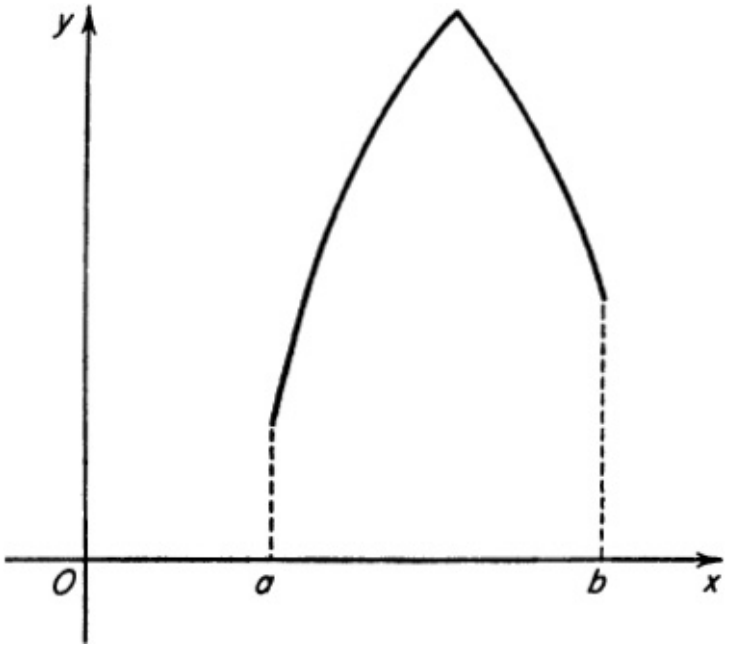


Figura 12.

$$\frac{f(x+h) - f(x)}{h}$$

poiché $h \to 0$ non si avvicina a un limite finito. Questo caso ovviamente si verifica in ogni punto di discontinuità della funzione $f(x)$, poiché qui il rapporto

$$\frac{f(x+h) - f(x)}{h} \qquad (10)$$

ha un numeratore che non si avvicina allo zero mentre il denominatore diminuisce senza limiti. La derivata potrebbe anche non esistere nel punto in cui la funzione è continua. Un semplice esempio è dato da un punto qualsiasi in cui il grafico della funzione forma un angolo

(figura 12). In tale punto la curva del grafico non ha una tangente definita e di conseguenza la funzione non ha derivata. Spesso in tali punti l'espressione (10) si avvicina a valori diversi, a seconda che h si avvicini allo zero da destra o da sinistra, così che se h si avvicina allo zero in modo arbitrario, il rapporto (10) semplicemente non ha limiti. Un esempio di una funzione più complicata senza derivata è dato da

$$y = \begin{cases} x \sin \frac{1}{x} & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

Il grafico di questa funzione è disegnato in figura 13. Nel punto $x = 0$ non ha derivata perché, come è evidente dal grafico, la secante OA non si avvicina ad alcuna posizione definita anche quando $A \to 0$ da un lato. Infatti la secante OA oscilla all'infinito avanti e indietro tra la retta OM e la retta OL. Il corrispondente rapporto (10) in questo caso non ha limite, anche se h conserva lo stesso segno avvicinandosi allo zero.

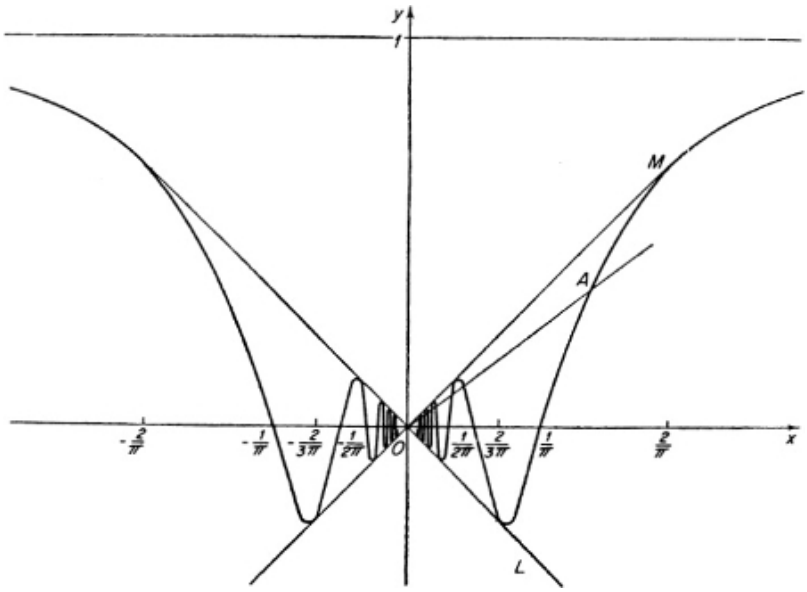


Figura 13.

Notiamo infine che è possibile definire, in modo puramente analitico mediante una formula, una funzione continua che non ha derivata in nessun punto. Un esempio di tale funzione fu fornito per la prima volta dall'eminente matematico tedesco del secolo scorso, Weierstrass.

Di conseguenza la classe delle funzioni derivabili è notevolmente più ristretta di quella delle funzioni continue.

Passiamo ora al calcolo vero e proprio delle derivate delle funzioni più semplici.

1. $y = c$, dove c è una costante. Una costante può essere considerata come un caso speciale di una funzione che rimane uguale allo stesso numero per x arbitrario. Il suo grafico è una linea retta parallela all'asse x ad una distanza pari a c. Questa retta forma con l'asse x un angolo $\alpha = 0$, e ovviamente la derivata di una costante è identicamente uguale a zero: $y' = (c)' = 0$. Dal

punto di vista della meccanica, questa equazione significa che la velocità di un punto fisso è uguale a zero.

2. $y = x^2$

$$\frac{f(x+h)-f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = 2x + h.$$

Poiché $h \to 0$ si ottiene [12] nel limite $2x$; di conseguenza

$$y' = (x^2)' = 2x.$$

3. $y = x^n$ (n un numero intero positivo).

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &= \frac{(x+h)^n - x^n}{h} \\ &= \frac{x^n + nx^{n-1}h + \frac{n(n-1)}{2!}x^{n-2}h^2 + \cdots + h^n - x^n}{h} \\ &= nx^{n-1} + \frac{n(n-1)}{2!}x^{n-2}h + \cdots + h^{n-1}.\end{aligned}$$

Ogni addendo a destra, a cominciare dal secondo, si avvicina allo zero come $h \to 0$; di conseguenza

$$y' = (x^n)' = nx^{n-1}.$$

Questa formula rimane vera per n arbitrari positivi o negativi, frazionari o addirittura irrazionali, anche se la dimostrazione deve quindi essere diversa. Utilizzeremo questo fatto senza dimostrarlo. Così ad esempio

$$(\sqrt{x})' = (x^{\frac{1}{2}})' = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}, (x > 0);$$

$$(\sqrt[3]{x})' = (x^{\frac{1}{3}})' = \frac{1}{3}x^{-\frac{2}{3}} = \frac{1}{3\sqrt[3]{x^2}}, (x \neq 0);$$

$$\left(\frac{1}{x}\right)' = (x^{-1})' = -1 \cdot x^{-2} = -\frac{1}{x^2}, (x \neq 0);$$

$$(x^\pi)' = \pi x^{\pi-1}, \qquad (x > 0).$$

4. $y =$ peccato x.

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

Come spiegato in precedenza, la prima frazione si avvicina all'unità come $h \to 0$, e il cos $(x + h/2)$ ovviamente si avvicina al cos x. Quindi la derivata del seno è uguale al coseno

$$y' = (\sin x)' = \cos x.$$

Suggeriamo al lettore di dimostrarlo con lo stesso tipo di argomento

$$(\cos x)' = -\sin x.$$

5. In precedenza (Capitolo II, §3) abbiamo già notato l'esistenza del limite

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e = 2.71828 \cdots .$$

Abbiamo inoltre osservato che per il calcolo di tale limite non gioca un ruolo essenziale il fatto che n assume solo valori integrali positivi. È importante solo che l'infinitesimale $1/n$, che viene sommato all'unità, e l'esponente n, che aumenta oltre ogni limite, siano reciproci tra loro.

Utilizzando questa affermazione possiamo facilmente trovare la derivata del logaritmo $y = \log_a x$

$$\frac{\log_a (x + h) - \log_a x}{h} = \frac{1}{h} \log_a \frac{x + h}{x} = \frac{1}{x} \log_a \left(1 + \frac{h}{x}\right)^{x/h}.$$

La continuità del logaritmo permette di sostituire la quantità sotto il segno logaritmico con il suo limite, che è pari a e; così

$$\lim_{h\to 0} \left(1 + \frac{h}{x}\right)^{x/h} = e$$

(in questo caso il ruolo di $n \to \infty$ è giocato dalla quantità crescente x/h). Di conseguenza, otteniamo la regola per derivare un logaritmo

$$(\log_a x)' = \frac{1}{x} \log_a e.$$

Questa regola diventa particolarmente semplice se come base dei nostri logaritmi scegliamo il numero e. I logaritmi in questa base sono detti logaritmi naturali e sono indicati con $\ln x$. Possiamo scrivere

$$(\log_e x)' = \frac{1}{x}$$

o ancora

$$(\ln x)' = \frac{1}{x}.$$

§ 6. Regole di derivazione

Dagli esempi forniti in precedenza può sembrare che il calcolo della derivata di ogni nuova funzione richieda l'invenzione di nuovi metodi. Non è questo il caso. Lo sviluppo dell'analisi è stato reso possibile non poco dalla scoperta di un semplice metodo unificato per trovare la derivata di una funzione arbitraria "elementare" (cioè una funzione che può essere espressa da una formula costituita da una combinazione finita delle operazioni algebriche fondamentali, delle funzioni trigonometriche, dell'operazione di elevazione a una potenza e dell'estrazione dei logaritmi). Alla base di questo metodo ci sono le cosiddette regole di derivazione. Esse consistono da una serie di teoremi che ci permettono di ridurre i problemi più complicati a quelli più semplici.

Spiegheremo qui le regole di derivazione e cercheremo di essere molto brevi nel dedurle. Se il lettore desidera formarsi semplicemente un'idea generale dell'analisi, può omettere la presente sezione, ricordando solo che esiste un mezzo per trovare effettivamente la derivata di qualsiasi funzione elementare. In questo caso sarà ovviamente necessario che egli si fidi di alcuni calcoli dei nostri esempi successivi.

Derivata di una somma. Supponiamo che y sia dato in funzione di x dall'espressione

$$y = \phi(x) + \psi(x),$$

dove $u = \phi(x)$ e $v = \psi(x)$ sono funzioni note di x. Assumiamo inoltre di poter trovare le derivate delle funzioni u e v. Come possiamo allora trovare la derivata della funzione y? La risposta è semplice

$$y' = (u + v)' = u' + v'. \qquad (11)$$

Diamo infatti a x un incremento Δx; quindi u, v e y riceveranno ciascuno un incremento Δu, Δv e Δy, collegati dall'equazione

$$\Delta y = \Delta u + \Delta v.$$

Così[13]

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x},$$

e dopo il passaggio al limite per $\Delta x \to 0$ si ottiene subito la formula (11), se ovviamente le funzioni u e v hanno derivate.

Analogamente possiamo derivare la formula per derivare la differenza di due funzioni

$$(u - v)' = u' - v'. \qquad (12)$$

Derivato di un prodotto. La regola per la derivazione di un prodotto è un po' più complicata. Esiste la derivata del prodotto di due funzioni, ciascuna delle quali ha una derivata, ed è

uguale alla somma del prodotto della prima funzione per la derivata della seconda e del prodotto della seconda per la derivata della prima; questo è

$$(uv)' = uv' + vu'. \qquad (13)$$

Diamo infatti a x un incremento Δx. Quindi le funzioni u, v e y = uv riceveranno gli incrementi ϕu, Δv, Δy, soddisfacendo la relazione

$$\Delta y = (u + \Delta u)(v + \Delta v) - uv = u\,\Delta v + v\,\Delta u + \Delta u\,\Delta v,$$

da cui

$$\frac{\Delta y}{\Delta x} = u\frac{\Delta v}{\Delta x} + v\frac{\Delta u}{\Delta x} + \Delta u\frac{\Delta v}{\Delta x}.$$

Dopo il passaggio al limite per $\Delta x \to 0$ i primi due addendi a destra producono il lato destro della formula (13) mentre il terzo addendi svanisce.[14] Di conseguenza, nel limite si ottiene la regola (13).

Nel caso particolare $v = c$ = costante, abbiamo

$$(cu)' = cu' + uc' = cu', \qquad (14)$$

poiché la derivata di una costante è uguale a zero.

Derivata di un quoziente. Sia $y = u/v$, dove u e v hanno una derivata per un dato x, con $v \equiv 0$ per quel valore di x. Ovviamente

$$\Delta y = \frac{u + \Delta u}{v + \Delta v} - \frac{u}{v} = \frac{v\,\Delta u - u\,\Delta v}{(v + \Delta v)\,v},$$

da cui

$$\frac{\Delta y}{\Delta x} = \frac{v\dfrac{\Delta y}{\Delta x} - u\dfrac{\Delta v}{\Delta x}}{(v + \Delta v)\,v} \to \frac{vu' - uv'}{v^2}\ (\Delta x \to 0).$$

Anche qui ci siamo serviti del fatto che per una funzione v che ha una derivata abbiamo necessariamente $\Delta v \to 0$, quando $\Delta x \to 0$. Quindi

$$\left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2}. \qquad (15)$$

Diamo alcuni esempi di applicazione di queste regole

$$(2x^3 - 5)' = 2(x^3)' - (5)' = 2 \cdot 3x^2 - 0 = 6x^2;$$

$$(x^2 \sin x)' = x^2(\sin x)' + (x^2)' \sin x = x^2 \cos x + 2x \sin x;$$

$$(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos x(\sin x)' - \sin x(\cos x)'}{\cos^2 x}$$

$$= \frac{\cos x \cdot \cos x - \sin x(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x} = \sec^2 x.$$

Consigliamo al lettore di provare da solo la formula

$$(\cot x)' = -\csc^2 x.$$

Derivata della funzione inversa. Consideriamo una funzione $y = f(x)$, che è continua e crescente (decrescente) sull'intervallo $[a, b]$. Per crescente (decrescente) si intende che ad un valore maggiore di x nell'intervallo $[a, b]$ corrisponde un valore maggiore (minore) di y (figura 14).

Sia $c = f(a)$ e $d = f(b)$. Nella figura 14 è evidente che ad ogni valore di y dell'intervallo $[c, d]$ (o $[d, c]$, rispettivamente) corrisponde esattamente un valore di x dell'intervallo $[a, b]$ tale che $y = f(x)$. Quindi sull'intervallo $[c, d]$ (o $[d, c]$) abbiamo una funzione completamente determinata $x = \phi(y)$, che è chiamata funzione inversa di $y = f(x)$. Nella figura 14 è chiaro che la funzione $\phi(y)$ è continua, un fatto dimostrato nell'analisi moderna con metodi strettamente analitici. Ora lasciamo che Δx e Δy corrispondano rispettivamente agli incrementi in x e y. È evidente che

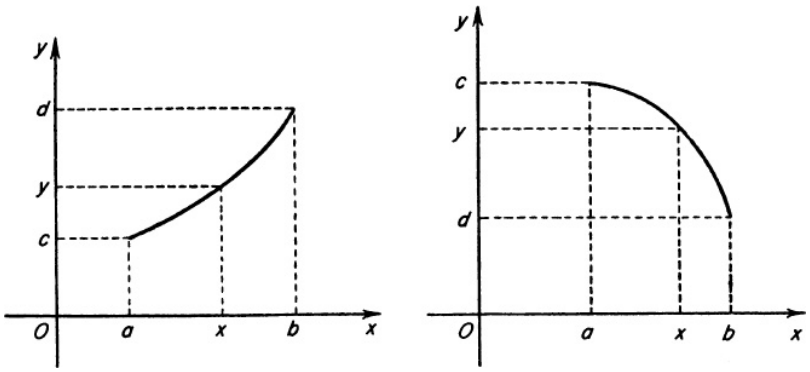


Figura 14.

$$\frac{\Delta y}{\Delta x} = \frac{1}{\Delta x/\Delta y}, \text{ if } \Delta y \neq 0.$$

Al limite questo ci dà una semplice relazione tra le derivate delle funzioni dirette e inverse

$$y'_x = \frac{1}{x'_y}. \qquad (16)$$

Utilizziamo questa relazione per trovare la derivata della funzione $y = a^x$. La funzione inversa è $x = \log_a y$, che siamo già in grado di derivare, e quindi possiamo scrivere

$$(a^x)'_x = \frac{1}{(\log_a y)'_y} = \frac{1}{1/y\,(\log_a e)} = y \log_e a = a^x \ln a. \qquad (17)$$

In particolare $(e^x)' = e^x$.

Come altro esempio prendiamo $y = \arcsin x$. La funzione inversa è $x = \sin y$. Così

$$(\arcsin x)'_x = \frac{1}{(\sin y)'_y} = \frac{1}{\cos y} = \frac{1}{\sqrt{1 - (\sin y)^2}} = \frac{1}{\sqrt{1 - x^2}}.$$

Tabella dei derivati. Tabuliamo le derivate delle funzioni elementari più semplici (Tabella 2).

Tabella 2

y	y'	y	y'	y	y'
c	0	$\ln x$	$\frac{1}{x}$	$\tan x$	$\sec^2 x$
x^a	ax^{a-1}	$\log_a x$	$\frac{1}{x}\log_a e$	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
e^x	e^x	$\sin x$	$\cos x$	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
a^x	$a^x \ln a$	$\cos x$	$-\sin x$	$\arctan x$	$\frac{1}{1+x^2}$

Queste formule sono state calcolate e spiegate in precedenza, ad eccezione delle ultime due che il lettore può, se lo desidera, ricavare facilmente da solo utilizzando la regola per la derivazione di una funzione inversa.

Calcolo della derivata di una funzione di una funzione. Resta da considerare l'ultima e più difficile regola per la derivazione. Il lettore in possesso di questa regola e di una serie di tabelle può con pieno diritto ritenere di essere in grado di derivare qualsiasi funzione elementare.

Per applicare la regola che stiamo per dare è necessario avere ben chiaro come è costruita la funzione che si vuole derivare; cioè quali operazioni dobbiamo eseguire sulla variabile indipendente x, e in quale ordine, per produrre il valore della variabile dipendente y.

Ad esempio, per calcolare la funzione

$$y = \sin x^2,$$

è necessario innanzitutto elevare x alla seconda potenza e poi prendere il seno della grandezza così ottenuta, procedimento che può essere descritto nel seguente modo: $y = \sin u$, dove $u = x^2$.

D'altra parte, per calcolare la funzione

$$y = \sin^2 x,$$

è necessario innanzitutto trovare il seno di x, e poi elevare alla seconda potenza il valore così trovato, procedura che può essere scritta così: $y = u^2$, dove $u = \sin x$.

Ecco alcuni esempi:

1. $y = (3x + 4)^3$, $y = u^3$, $u = 3x + 4$.

2. $y = \sqrt{1 - x^2}$, $y = u^{\frac{1}{2}}$, $u = 1 - x^2$.

3. $y = e^{kx}$; $y = e^u$, $u = kx$.

Nei casi più complicati abbiamo una catena di relazioni semplici, che può avere più anelli. Per esempio,

4. $y = \cos^3 x^2$; $y = u^3$; $v = \cos v = x^2$.

Se y è una funzione della variabile u

$$y = f(u), \tag{18}$$

e u a sua volta è una funzione della variabile x

$$u = \phi(x), \tag{19}$$

quindi y, essendo una funzione di u, è anche una certa funzione di x, che può essere indicata come segue

$$y = F(x) = f[\phi(x)]. \tag{20}$$

Considerando casi più complicati possiamo formare, ad esempio, la funzione

$$y = \Phi(x) = f\{\phi[\psi(x)]\},$$

che è equivalente alle equazioni

$$y = f(u), u = \phi(v), v = \psi(x),$$

e potremmo formare catene ancora più lunghe.

Mostriamo ora come calcolare la derivata della funzione $F(x)$ definita dall'equazione (20) se conosciamo la derivata di $f(u)$ rispetto a u e la derivata di $\phi(x)$ rispetto a x.

Diamo a x l'incremento Δx; quindi per (19) u riceverà un certo incremento Δu e per (18) y riceverà un incremento Δy. Così possiamo scrivere

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}.$$

Ora lasciamo che Δx si avvicini allo zero. Quindi $\Delta u/\Delta x \to u'_x$. Inoltre, dalla continuità di u, l'incremento $\Delta u \to 0$, e quindi $\Delta y/\Delta u \to y'_u$ (si è ipotizzata l'esistenza delle derivate y'_u e u'_x).

Così abbiamo dimostrato l'importante formula per la derivata di una funzione di una funzione[15]

$$y'_x = y'_u u'_x . \qquad (21)$$

Calcoliamo, dalla formula (21) e dalla tabella fondamentale delle derivate data, le derivate delle funzioni che abbiamo considerato:

1. $y = (3x + 4)^3 = u^3$, $y'_x = (u^3)'_u(3x + 4)'_x = 3u^2 \cdot 3 = 9(3x + 4)^2 = 3u^2 \cdot 3 = 9(3x + 4)^2$.

2.

$$y = \sqrt{1 - x^2} = u^{\frac{1}{2}},\ y'_x = (u^{\frac{1}{2}})'_x\,(1 - x^2)'_x = \frac{1}{2}u^{-\frac{1}{2}}(-2x) = -\frac{x}{\sqrt{1 - x^2}}.$$

3. $y = e^{kx} = e^u$, $y'_x = (e^u)'_u \cdot u'_x = e^u \cdot k = ke^{kx}$.

Se $y = f(u)$, $u = \phi(v)$, $v = f(x)$, allora

$$y'_x = y'_u \cdot u_x = y'_u(u'_v \cdot v'_x) = y'_u \cdot u'_v \cdot v'_x .$$

È chiaro come generalizzare questa formula per il caso di un numero arbitrario (finito) di funzioni nella catena. Per esempio,

4.

$$y = \cos^3 x^2;\ y'_x = (u^3)'_u\,(\cos v)'_v \cdot (x^2)'_x = 3u^2(-\sin v) \cdot 2x = -6x\cos^2 x^2 \sin x^2.$$

Nella nostra spiegazione su come calcolare la derivata di una funzione, abbiamo introdotto le variabili intermedie u, v, . . . Ma in realtà, dopo un po' di pratica, se ne può fare a meno, semplicemente tenendo presenti le funzioni che denotano.

Le funzioni elementari. Per chiudere la presente sezione ricordiamo che le funzioni le cui derivate sono state elencate in forma tabellare (Tabella 2) possono essere utilizzate per definire le cosiddette funzioni elementari. Queste funzioni elementari sono definite come quelle funzioni che possono essere ottenute dalle funzioni semplici precedenti mediante le quattro operazioni aritmetiche e l'operazione di prendere una funzione di una funzione, ciascuna di queste operazioni essendo eseguita un numero finito di volte.

Ad esempio, il polinomio $x^2 - 2x^2 + 3x - 5$ è una funzione elementare poiché è ottenuta mediante operazioni aritmetiche da un numero di funzioni della forma x^k. Anche la funzione $\ln \sqrt{1 - x^2}$ è elementare, poiché si ottiene dal polinomio $u = 1 - x^2$ mediante l'operazione $v = u^{1/2}$, e successivamente l'operazione $\ln v$.

Le regole di derivazione discusse in precedenza sono sufficienti per ottenere la derivata di qualsiasi funzione elementare, non appena conosciamo le derivate delle funzioni elementari più semplici.

§ 7. Massimi e minimi; studio dei grafici delle funzioni

Una delle applicazioni più semplici e importanti della derivata è nella teoria dei massimi e dei minimi. Supponiamo che su un certo intervallo $a \leqslant x \leqslant b$ ci sia data una funzione $y = f(x)$ che non solo è continua ma ha anche una derivata in ogni punto. La nostra capacità di calcolare la derivata ci consente di formare un quadro chiaro del grafico della funzione. Su un intervallo in cui la derivata è sempre positiva la tangente al grafico sarà diretta verso l'alto. In tale intervallo la funzione aumenterà; cioè ad un valore maggiore di x corrisponderà un valore maggiore di $f(x)$. D'altra parte, su un intervallo in cui la derivata è sempre negativa, la funzione diminuirà; il grafico scorrerà verso il basso.

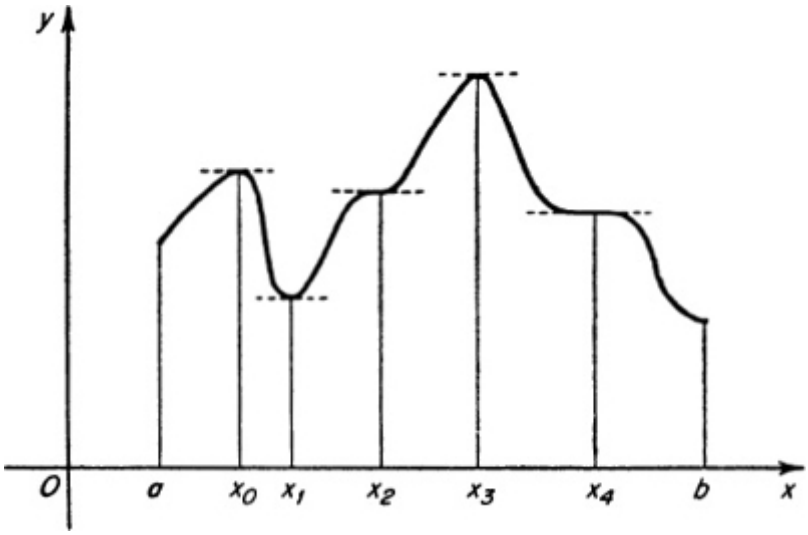


Figura 15.

Massimo e minimo. In figura 15 abbiamo disegnato il grafico di una funzione $y = f(x)$ definita sull'intervallo (a, b). Di particolare interesse sono i punti di questo grafico le cui ascisse sono x_0, x_1, x_3.

Nel punto x_0 si dice che la funzione $f(x)$ ha un massimo locale; con questo intendiamo che in questo punto $f(x)$ è maggiore che nei punti vicini; più precisamente $f(x_0) \geqslant f(x)$ per ogni x in un certo intervallo attorno al punto x_0.

Un minimo locale è definito in modo analogo. Per la nostra funzione si verifica un massimo locale nei punti x_0 e x_3, e un minimo locale nel punto x_1.

In ogni punto di massimo o di minimo, se è interno all'intervallo $[a, b]$, cioè se non coincide con uno dei punti finali a o b, la derivata deve essere uguale a zero.

Quest'ultima affermazione, molto importante, consegue immediatamente dalla definizione di derivata come limite del rapporto $\Delta y/\Delta x$. Infatti, se ci spostiamo a breve distanza dal punto massimo, allora $\Delta y \leqslant 0$. Pertanto per Δx positivo il rapporto $\Delta y/\Delta x$ non è positivo, mentre per Δx negativo il rapporto $\Delta y/\Delta x$ non è negativo. Il limite di questo rapporto, che esiste per

ipotesi, non può quindi essere né positivo né negativo e resta solo la possibilità che sia zero. Osservando il diagramma si vede che ciò significa che nei punti di massimo o di minimo (è consuetudine tralasciare la parola “locale”, anche se è inteso) la tangente al grafico è orizzontale. Nella figura 15 si nota che nei punti x_2, e x_4 anche la tangente è orizzontale, così come lo è nei punti x_0, x_1, x_3, anche se in questi punti la funzione non ha né massimo né minimo. In generale, possono esserci più punti in cui la derivata della funzione è uguale a zero (punti stazionari) rispetto ai punti di massimo o di minimo.

Determinazione dei valori massimo e minimo di una funzione. In numerose questioni tecniche è necessario trovare il punto x in cui una determinata funzione $f(x)$ raggiunge il suo valore massimo o minimo in un dato intervallo.

Nel caso in cui siamo interessati al valore maggiore, dobbiamo trovare x_0 sull'intervallo $[a, b]$ per il quale tra tutti x su $[a, b]$ è soddisfatta la disuguaglianza $f(x_0) \geqslant f(x)$.

Ma ora sorge la questione fondamentale: se esista in generale un punto del genere. Con i metodi dell'analisi moderna è possibile dimostrare il seguente teorema di esistenza: Se la funzione $f(x)$ è continua su un intervallo finito, allora esiste almeno un punto sull'intervallo per il quale la funzione raggiunge il suo valore massimo (minimo) sull'intervallo $[a, b]$.

Da quanto già detto ne consegue che questi punti massimi o minimi vanno ricercati tra i punti “stazionari”. Questo fatto è la base per il seguente metodo ben noto per trovare i massimi e i minimi.

Per prima cosa troviamo la derivata di, $f(x)$ e poi risolviamo l'equazione ottenuta ponendola uguale a zero

$$f'(x) = 0.$$

Se $x_1, x_2, \cdots, x_n$, sono le radici di questa equazione, confrontiamo quindi i numeri $f(x_1), f(x_2), \cdots, f(x_n)$ tra loro Naturalmente bisogna tenere conto che il massimo o il minimo della funzione possono trovarsi non all'interno dell'intervallo ma alla fine (come nel caso del minimo in figura 15) o in un punto in cui la funzione non ha derivata (come in figura 15). figura 12). Quindi ai punti $x_1, x_2, \cdots, x_n$, dobbiamo aggiungere gli estremi a e b dell'intervallo ed anche quei punti, se esistono, in cui non esiste derivata la funzione in tutti questi punti e scegliere tra essi il maggiore o il minore.

Rispetto all'enunciato teorema di esistenza, è importante aggiungere che tale teorema cessa, in generale, di valere nel caso in cui la funzione $f(x)$ sia continua solo sull'intervallo (a, b); cioè sull'insieme dei punti x che soddisfano le disuguaglianze $a < x < b$. Lasciamo al lettore la considerazione del fatto che la funzione $1/x$ non ha né massimo né minimo sull'intervallo (0, 1).

Vediamo alcuni esempi.

Da un pezzo quadrato di stagno di lato a è necessario realizzare una scatola rettangolare aperta di massimo volume. Se dagli angoli del quadrato originale togliamo i quadrati di lato x (vedi §2, esempio 2) otteniamo una scatola con il volume

$$V = x(a - 2x)^2.$$

Il nostro problema diventa quindi trovare il valore di x per il quale la funzione $V(x)$ raggiunge il suo valore massimo nell'intervallo $0 \leqslant x \leqslant a/2$. Secondo la regola, troviamo la derivata e la poniamo uguale a zero

$$V'(x) = (a - 2x)^2 - 4x(a - 2x) = 0.$$

Risolvendo questa equazione, troviamo le due radici

$$x_1 = \frac{a}{2}, x_2 = \frac{a}{6}.$$

A questi aggiungiamo l'estremità sinistra dell'intervallo (l'estremità destra è identica a x_1) e confrontiamo i valori della funzione in questi punti

$$V(0) = 0; V\left(\frac{a}{6}\right) = \frac{2}{27}a^3, V\left(\frac{a}{2}\right) = 0.$$

Pertanto la scatola avrà il volume maggiore, pari a $2/27\ a^3$, per l'altezza $x = a/6$.

Come secondo esempio, esaminiamo il problema della lampada della pista di pattinaggio (vedi §2, esempio 3). A quale altezza h dobbiamo posizionare la lampada affinché il bordo della pista riceva la massima illuminazione?

Per la formula (3) §2, il nostro problema si riduce a determinare il valore di h per $T = A \sin \alpha/h^2 + r^2$ assume il suo valore massimo. Invece di h qui è più conveniente trovare l'angolo α (figura 3, Capitolo I). Abbiamo

$$h = r \tan \alpha,$$

affinché

$$T = \frac{A}{r^2}\frac{\sin \alpha}{1 + \tan^2 \alpha} = \frac{A}{r^2}\sin \alpha \cos^2 \alpha.$$

Quindi occorre trovare il massimo della funzione $T(\alpha)$ tra quei valori di α che soddisfano la disuguaglianza $0 < \alpha < \pi 2$. Per fare ciò, troviamo la derivata e la poniamo uguale a zero

$$T'(\alpha) = \frac{A}{r^2}(\cos^3 \alpha - 2 \sin^2 \alpha \cos \alpha) = 0.$$

Questa equazione si divide in due

$$\cos \alpha = 0, \cos^2 \alpha - 2 \sin^2 \alpha = 0.$$

La prima equazione ha radice $\alpha = \pi 2$, che coincide con la fine dell'intervallo $(0, \pi/2)$. La seconda equazione può essere messa nella forma

$$\tan^2 \alpha = \tfrac{1}{2}.$$

Ma poiché $0 < \alpha < \pi/2$, abbiamo il risultato $\alpha \approx 35°15'$. Questo è quindi il valore per il quale la funzione $T(\alpha)$ raggiunge il suo massimo (alle estremità dell'intervallo $T = 0$). L'altezza desiderata h è quindi pari a

$$h = r \tan \alpha = \frac{r}{\sqrt{2}} \approx 0.7r.$$

Per una migliore illuminazione del bordo pista la lampada dovrà essere posizionata ad un'altezza pari a circa 0,7 volte il raggio.

Ma ora supponiamo che le strutture a nostra disposizione non ci permettano di sollevare la lampada ad un'altezza maggiore di un certo H. Quindi l'angolo a può variare non da 0 a $\pi/2$ ma solo entro i limiti più ristretti $0 < \alpha \leqslant \text{arc tan } (H/r)$. Ad esempio, supponiamo che $r = 12$ metri e $H = 9$ metri. In questo caso è infatti possibile alzare la lampada all'altezza $h = r/\sqrt{2}$, che equivale a poco più di 8 metri, quindi è quello che dobbiamo fare. Ma se H è inferiore a 8 metri (ad esempio, se abbiamo a nostra disposizione solo un palo lungo 6 metri), allora risulta che la derivata della funzione $T(\alpha)$ nell'intervallo [O, arc tan (H/r)] non è mai uguale a zero. In questo caso il massimo si raggiunge alla fine dell'intervallo e la lampada dovrà essere sollevata alla massima altezza possibile $H = 6$ metri.

Finora abbiamo considerato una funzione su un intervallo finito. Se l'intervallo è di lunghezza infinita, anche una funzione continua potrebbe non riuscire a raggiungere il suo valore massimo o minimo ma potrebbe, ad esempio, continuare a crescere o diminuire man mano che x si avvicina all'infinito.

Pertanto le funzioni $y = kx + b$ (vedi figura 5, §2 di questo capitolo), $y = \text{arc tan } x$ (figura 16a), $y = \ln x$ (figura 16b) non raggiungono da nessuna parte né un massimo né un minimo. La funzione $y = e^{-x^2}$ (figura 16c) raggiunge il suo massimo nel punto $x = 0$, ma non raggiunge da nessuna parte il minimo. Per quanto riguarda la funzione $y = x/(1 + x^2)$ (figura 16d), essa raggiunge il suo minimo nel punto $x = -1$ e il suo massimo nel punto $x = 1$.[E1]

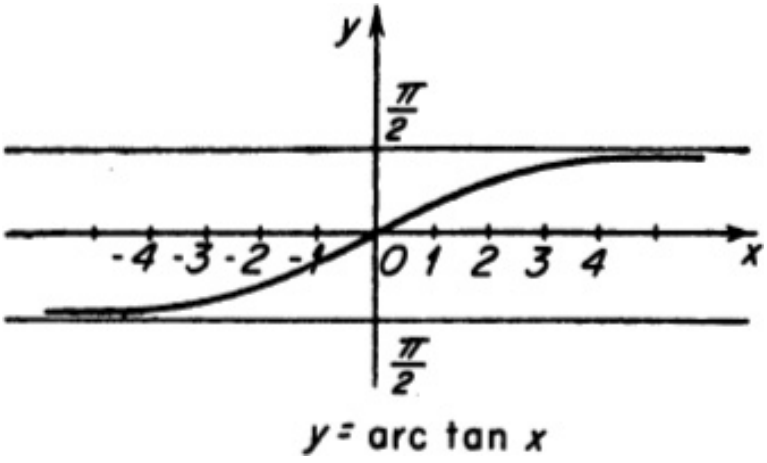


Figura 16a.

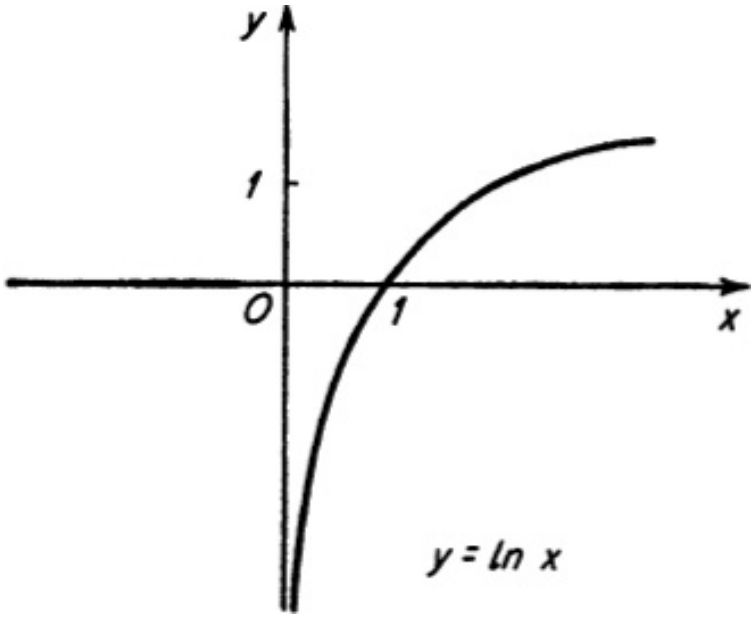


Figura 16b.

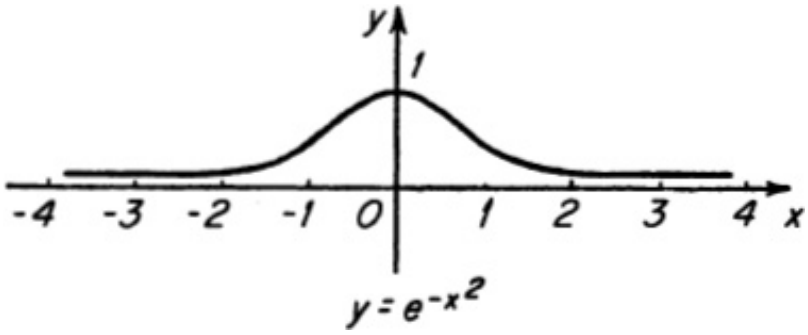


Figura 16c.

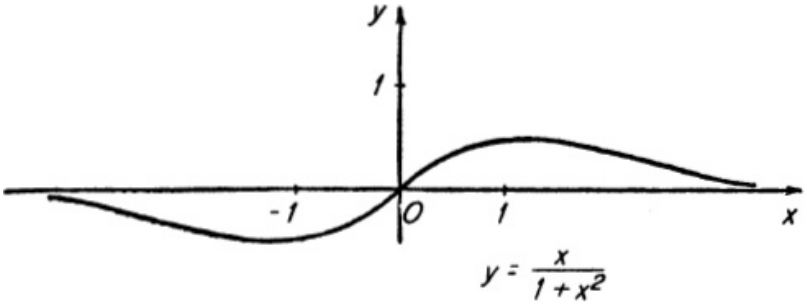


Figura 16d.

Nel caso di un intervallo di durata infinita l'indagine può ridursi alle regole ordinarie. È solo necessario considerare al posto di $f(a)$ e $f(b)$ i limiti

$$A = \lim_{x \to -\infty} f(x), \quad B = \lim_{x \to +\infty} f(x).$$

Derivati di ordine superiore. Abbiamo appena visto come, per studiare più da vicino il grafico di una funzione, dobbiamo esaminare le variazioni della sua derivata $f'(x)$. Questa derivata è una funzione di x, quindi possiamo a nostra volta trovare la sua derivata.

La derivata della derivata è detta derivata seconda e si indica con

$$[y']' = y'' \text{ or } [f'(x)]' = f''(x).$$

Analogamente possiamo calcolare la derivata terza

$$[y'']' = y''' \text{ or } [f''(x)]' = f'''(x)$$

e più in generale la derivata *n*esimo o, come viene anche chiamata, la derivata dell'ordine *n*esimo

$$y^{(n)} = f^{(n)}(x).$$

Naturalmente bisogna tenere presente che, per un certo valore di x (o anche per tutti i valori di x) questa successione può interrompersi in corrispondenza della derivata di qualche ordine, ad esempio k; può succedere che esista $f^{(k)}(x)$ ma non $f^{(k+1)}(x)$. Le derivate di ordine arbitrario appariranno più avanti nel §9 in relazione alla formula di Taylor. Per il momento ci limitiamo alla derivata seconda.

Significato della derivata seconda; convessità e concavità. La derivata seconda ha un significato semplice in meccanica. Sia $s = f(t)$ una legge del moto lungo una retta; allora s' è la velocità e s'' è la "velocità del cambiamento di velocità" o più semplicemente l'"accelerazione" del punto all'istante t. Ad esempio, per un corpo che cade sotto la forza di gravità

$$\begin{aligned} s &= \frac{gt^2}{2} + v_0 t + s_0 , \\ s' &= gt + v_0 , \\ s'' &= g, \end{aligned}$$

cioè l'accelerazione dei corpi in caduta è costante.

La derivata seconda ha anche un semplice significato geometrico. Proprio come il segno della derivata prima determina se la funzione è crescente o decrescente, così il segno della derivata seconda determina il lato verso il quale il grafico della funzione sarà curvo.

Supponiamo, ad esempio, che su un dato intervallo la derivata seconda sia ovunque positiva. Allora aumenta la derivata prima e quindi $f'(x) = \tan \alpha$ aumenta e aumenta l'angolo a di inclinazione della retta tangente stessa (figura 17). Pertanto, man mano che ci muoviamo lungo la curva, essa gira costantemente dalla stessa parte, cioè verso l'alto, ed è quindi, come si suol dire, "convessa verso il basso".

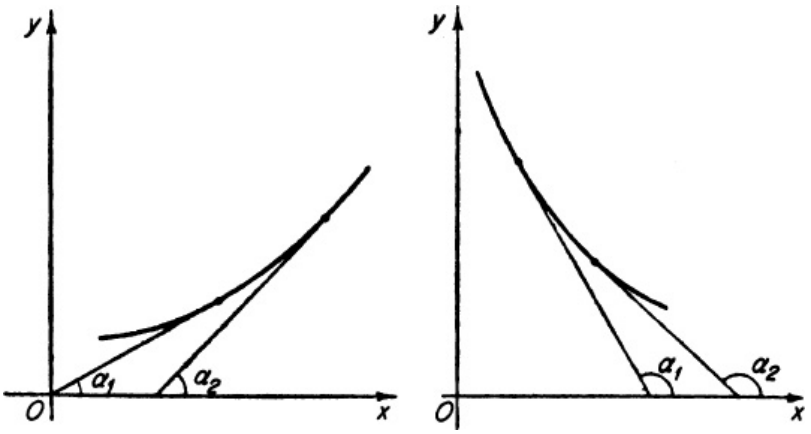


Figura 17.

D'altra parte, in una parte di curva dove la derivata seconda è costantemente negativa (figura 18) il grafico della funzione è "convesso verso l'alto".[16]

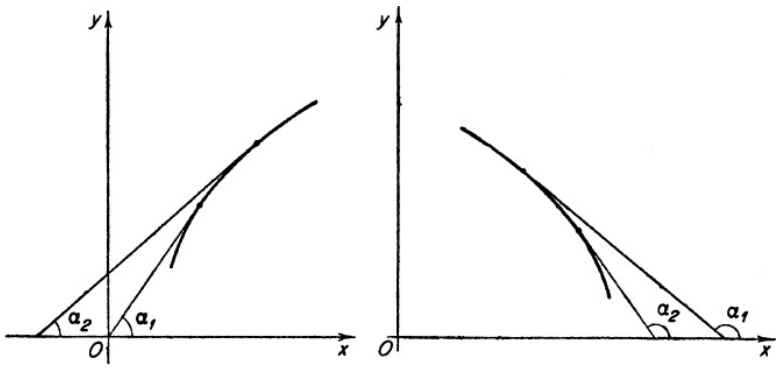


Figura 18.

Criteri per massimi e minimi; studio dei grafici delle curve. Se in tutto l'intervallo in cui varia x la curva è convessa verso l'alto e se in un certo punto x_0 di tale intervallo la derivata è uguale a zero, allora in questo punto la funzione raggiunge necessariamente il suo massimo; e il suo minimo nel caso di convessità verso il basso. Questa semplice considerazione spesso ci permette, dopo aver trovato un punto in cui la derivata è uguale a zero, di decidere quindi se in quel punto la funzione ha un massimo o un minimo locale.[17]

Esempio 1. Studiamo l'aspetto del grafico della funzione

$$f(x) = \frac{x^3}{3} - \frac{5x^2}{2} + 6x - 2.$$

Prendiamo la sua derivata prima e la poniamo uguale a zero,

$$f'(x) = x^2 - 5x + 6 = 0.$$

Le radici dell'equazione così ottenuta sono $x_1 = 2$, $x_2 = 3$. I corrispondenti valori della funzione sono

$$f(2) = 2\tfrac{2}{3},\ f(3) = 2\tfrac{1}{2}\,.$$

Segniamo quindi questi due punti sul diagramma. Insieme a questi possiamo anche segnare il punto con coordinate $x = 0$ e $y = f(0) = -2$ dove il grafico interseca l'asse y. La derivata seconda è $f''(x) = 2x - 5$. Questo si riduce a zero per $x = \frac{5}{2}$, quindi

$$f''(x) > 0 \text{ for } x > \tfrac{5}{2}\,,$$

$$f''(x) < 0 \text{ for } x < \tfrac{5}{2}\,.$$

Il punto

$$x = \frac{5}{2}\,,\ y = f\left(\frac{5}{2}\right) = 2\tfrac{7}{12}$$

è un punto di flesso del grafico. A sinistra di questo punto la curva è convessa verso l'alto, a destra è convessa verso il basso.

È ora evidente che il punto $x = 2$ è un punto di massimo e il punto $x = 3$ è un punto di minimo per la funzione.

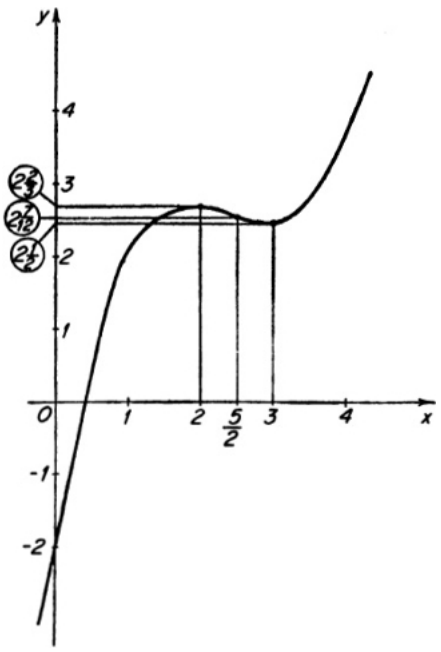


Figura 19.

Sulla base di questi risultati concludiamo che il grafico della funzione $y = f(x)$ ha l'aspetto schematizzato in figura 19. A destra del punto (0, −2) la curva sale al crescere di x, è convessa verso l'alto, e raggiunge il suo massimo nel punto $(2,\ 2\frac{2}{3})$, dopodiché comincia a cadere. Nel punto $(2\frac{1}{2},\ 2\frac{7}{12})$, dove $f''(x) = 0$, la convessità diventa concavità. Poi nel punto $(3,\ 2\frac{1}{2})$ la funzione raggiunge il suo minimo e da lì sale all'infinito. L'affermazione finale deriva dal fatto che il primo termine della funzione, quello contenente la potenza più alta (terza) di x, si avvicina all'infinito più velocemente del secondo e del terzo termine. Per lo stesso motivo il

grafico della funzione si avvicina a $-\infty$ poiché x assume valori negativi numericamente maggiori.

Esempio 2. Dimostreremo la disuguaglianza $e^x \geqslant 1 + x$ per x arbitrario. Consideriamo a questo scopo la funzione $f(x) = e^x - x - 1$. La sua derivata prima è $f'(x) = e^x - 1$, che si riduce a zero solo per $x = 0$. La derivata seconda $f''(x) = e^x > 0$ per tutti gli x. Di conseguenza il grafico della funzione $f(x)$ è convesso verso il basso. Il numero $f(0) = 0$ è il minimo per la funzione e $e^x - x - 1 \geqslant 0$ per tutti gli x.

Lo studio dei grafici ha molti scopi diversi. Spesso mostrano molto chiaramente, ad esempio, il numero di radici reali di una determinata equazione. Quindi, per dimostrare che l'equazione

$$xe^x = 2$$

ha un'unica radice reale, possiamo studiare i grafici delle funzioni $y = e^x$ e $y = 2/x$ (come schematizzato in figura 20). È facile vedere che questi grafici si intersecano in un solo punto, per cui l'equazione $e^x = 2/x$ ha esattamente una radice.

I metodi di analisi sono ampiamente applicati a problemi di calcolo approssimato delle radici di un'equazione. Su questo argomento si veda il Capitolo IV, §5.

§ 8. Incremento e differenziale di una funzione

Il differenziale di una funzione. Consideriamo una funzione $y = f(x)$ che ha una derivata. L'incremento di questa funzione

$$\Delta y = f(x + \Delta x) - f(x),$$

corrispondente all'incremento Δx, ha la proprietà che il rapporto $\Delta y/\Delta x$, poiché $\Delta x \to 0$, si avvicina ad un limite finito, pari alla derivata

$$\frac{\Delta y}{\Delta x} \to f'(x).$$

Questo fatto può essere scritto come un'uguaglianza

$$\frac{\Delta y}{\Delta x} = f'(x) + \alpha,$$

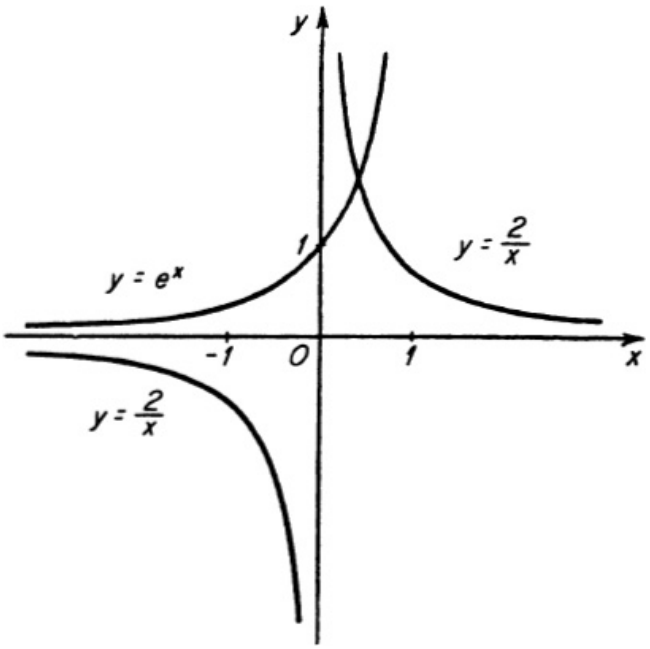


Figura 20.

dove il valore di a dipende da Δx in modo tale che poiché $\Delta x \to 0$, anche α si avvicina a zero. Pertanto l'incremento di una funzione può essere rappresentato nella forma

$$\Delta y = f'(x)\,\Delta x + \alpha \Delta x,$$

dove $\alpha \to 0$, se $\Delta x \to 0$.

Il primo addendo a destra di questa uguaglianza dipende da Δx in modo molto semplice, cioè è proporzionale a Δx. Si chiama differenziale della funzione, nel punto x corrispondente all'incremento dato Δx, ed è indicato con

$$dy = f'(x)\,\Delta x.$$

Il secondo addendo ha la proprietà caratteristica che, essendo $\Delta x \to 0$, si avvicina a zero più rapidamente di Δx, per effetto della presenza del fattore α. Si dice quindi che sia un infinitesimo di ordine superiore a Δx e, nel caso $f'(x) \neq 0$, è anche di ordine superiore al primo addendo. Con questo intendiamo che per Δx sufficientemente piccolo il secondo addendo è di per sé piccolo e anche il suo rapporto con Δx è arbitrariamente piccolo.

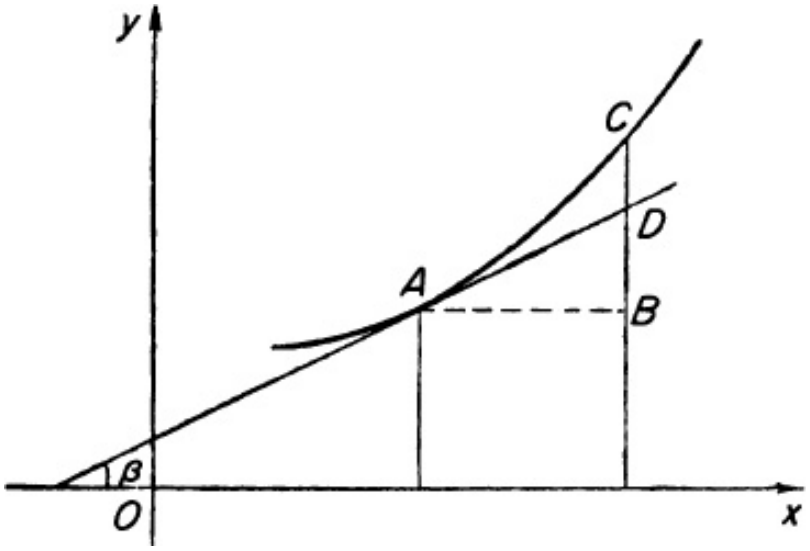


Figura 21.

La scomposizione di Δy in due addendi, di cui il primo (la parte principale) dipende linearmente da Δx e il secondo è trascurabile per Δx piccoli, può essere illustrata dalla figura 21. Il segmento BC = Δy, dove BC = BD + DC, BD = $\tan \beta \cdot \Delta x = f'(x)\,\Delta x = dy$ e DC è un infinitesimo di ordine superiore a Δx.

Nei problemi pratici il differenziale viene spesso utilizzato come valore approssimativo per l'incremento della funzione. Ad esempio, supponiamo di avere il problema di determinare il volume delle pareti di una scatola cubica chiusa le cui dimensioni interne sono 10 × 10 × 10 cm e lo spessore delle cui pareti è 0,05 cm. Se non è richiesta una grande precisione, possiamo argomentare come segue. Il volume di tutte le pareti della scatola rappresenta l'incremento Δy della funzione $y = x^3$ per $x = 10$ e $\Delta x = 0{,}1$. Quindi troviamo approssimativamente

$$\Delta y \approx dy = (x^3)'\,\Delta x = 3x^2\Delta x = 3 \cdot 10^2 \cdot 0.1 = 30\ \text{cm}^3.$$

Per simmetria nella notazione è consuetudine denotare l'incremento della variabile indipendente con dx e chiamarlo anche differenziale. Con questa notazione il differenziale della funzione può essere scritto così:

$$dy = f'(x)\,dx.$$

Quindi la derivata è il rapporto $f'(x) = dy/dx$ tra il differenziale della funzione e il differenziale della variabile indipendente.

Il differenziale di una funzione ha avuto origine storicamente nel concetto di “indivisibile”. Questo concetto, che dal punto di vista moderno non è mai stato definito con chiarezza, era a suo tempo, nel XVIII secolo, fondamentale nell'analisi matematica. Le idee che lo riguardano hanno subito cambiamenti essenziali nel corso di diversi secoli. L'indivisibile, e più tardi il differenziale di una funzione, furono rappresentati come veri e propri infinitesimi, come qualcosa di simile a una grandezza costante estremamente piccola, che però non era zero. La definizione fornita in questa sezione è quella accettata nell'analisi odierna. Secondo questa definizione il differenziale è una grandezza finita per ogni incremento Δx ed è allo stesso tempo proporzionale a Δx. L'altra proprietà fondamentale del differenziale, il carattere della sua differenza da Δy, può essere riconosciuta solo in movimento, per così dire: se consideriamo un incremento Δx che si avvicina allo zero (che è infinitesimale), allora la differenza tra dy e Δy sarà arbitrariamente piccola anche rispetto a Δx.

Questa sostituzione del differenziale al posto di piccoli incrementi della funzione costituisce la base della maggior parte delle applicazioni dell'analisi infinitesimale allo studio della natura. Il lettore lo vedrà in modo particolarmente chiaro nel caso delle equazioni differenziali, trattate in questo libro nei capitoli V e VI.

Pertanto, per determinare la funzione che rappresenta un dato processo fisico, proviamo innanzitutto a stabilire un'equazione che colleghi questa funzione in qualche modo definito con le sue derivate di vario ordine. Il metodo per ottenere tale equazione, chiamata equazione differenziale, spesso equivale a sostituire gli incrementi delle funzioni desiderate con i loro differenziali corrispondenti.

Ad esempio, risolviamo il seguente problema. In un sistema di coordinate rettangolare *Oxyz*, consideriamo la superficie ottenuta dalla rotazione della parabola la cui equazione (nel piano Oyz) è $z = y^2$. Questa superficie è chiamata paraboloide di rivoluzione (figura 22). Sia v il volume del corpo delimitato dal paraboloide e il piano parallelo al piano *Oxy* ad una distanza z da esso. È evidente che v è una funzione di z ($z > 0$).

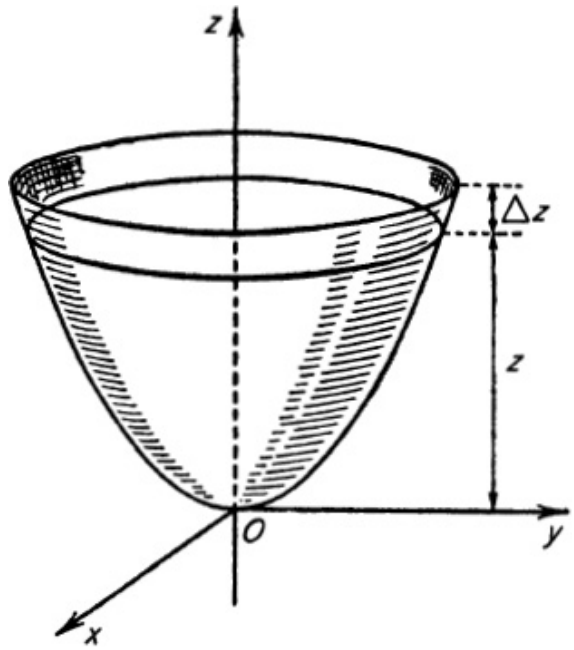


Figura 22.

Per determinare la funzione v, proviamo a trovare il suo differenziale dv. L'incremento Δv della funzione v nel punto z è pari al volume delimitato dal paraboloide e da due piani paralleli al piano Oxy alle distanze z e $z + \Delta z$ da esso.

È facile vedere che la grandezza di Δv è maggiore del volume del cilindro circolare di raggio $\sqrt{z}$ e altezza Δz ma inferiore a quello del cilindro circolare di raggio $\sqrt{z + \Delta z}$ e altezza Δz.

Così

$$\pi z\, \Delta z < \Delta v < \pi(z + \Delta z)\, \Delta z$$

e così

$$\Delta v = \pi(z + \theta\, \Delta z)\, \Delta z = \pi z\, \Delta z + \pi\theta\, \Delta z^2\,,$$

dove θ è un numero che dipende da Δz e che soddisfa la disuguaglianza $0 < \theta < 1$.

Siamo quindi riusciti a rappresentare l'incremento Δ_v sotto forma di una somma, il cui primo addendo è proporzionale a Δz, mentre il secondo è un infinitesimo di ordine superiore a Δz (in quanto $\Delta z \to 0$). Ne consegue che il primo addendo è il differenziale della funzione v

$$dv = \pi z\, \Delta z,$$

O

$$dv = \pi z\, dz,$$

poiché $\Delta z = dz$ per la variabile indipendente z.

L'equazione così ottenuta mette in relazione tra loro i differenziali dv e dz (delle variabili v e z) e viene quindi chiamata equazione differenziale.

Se teniamo conto di ciò

$$\frac{dv}{dz} = v',$$

dove v' è la derivata di v rispetto alla variabile z, la nostra equazione differenziale può essere scritta anche nella forma

$$v' = \pi z.$$

Per risolvere questa semplicissima equazione differenziale dobbiamo trovare una funzione di z la cui derivata sia uguale a πz. Problemi di questo genere sono trattati in maniera generale nei §§10 e 11, ma per il momento invitiamo il lettore a verificare che una soluzione della nostra equazione è data da $v = \pi z^2/2 + C$, dove per C possiamo scegliere un numero arbitrario.[18] Nel nostro caso il volume del corpo è ovviamente zero per $z = 0$ (vedi figura 22), per cui $C = 0$. Quindi la nostra funzione è data da $v = \pi z^2/2$.

Il teorema del valore medio ed esempi di applicazione. Il differenziale esprime il valore approssimativo dell'incremento della funzione in termini di incremento della variabile indipendente e della derivata nel punto iniziale. Quindi per l'incremento da $x = a$ a $x = b$, abbiamo

$$f(b) - f(a) \approx f'(a)(b - a).$$

È possibile ottenere un'equazione esatta di questo tipo se sostituiamo la derivata $f'(a)$ nel punto iniziale con la derivata in un punto intermedio, opportunamente scelto nell'intervallo (a, b). Più precisamente: Se $y = f(x)$ è una funzione derivabile sull'intervallo $a \leqslant x \leqslant b$, allora esiste un punto ξ, strettamente interno a tale intervallo, tale che vale la seguente uguaglianza esatta

$$f(b) - f(a) = f'(\xi)(b - a). \qquad (22)$$

L'interpretazione geometrica di questo "teorema del valore medio" (chiamato anche formula di Lagrange o formula delle differenze finite) è straordinariamente semplice. Siano A, B i punti sul grafico della funzione $f(x)$ che corrispondono a $x = a$ e $x = b$, e uniamo A e B dalla corda AB (figura 23). Spostiamo ora la retta AB, mantenendola costantemente parallela a se stessa, verso l'alto o verso il basso. Nel momento in cui questa linea retta taglia il grafico per l'ultima volta, sarà tangente al grafico in un certo punto C. A questo punto (sia l'ascissa corrispondente $x = \xi$), la retta tangente formerà lo stesso angolo di inclinazione α della corda AB. Ma per l'accordo abbiamo

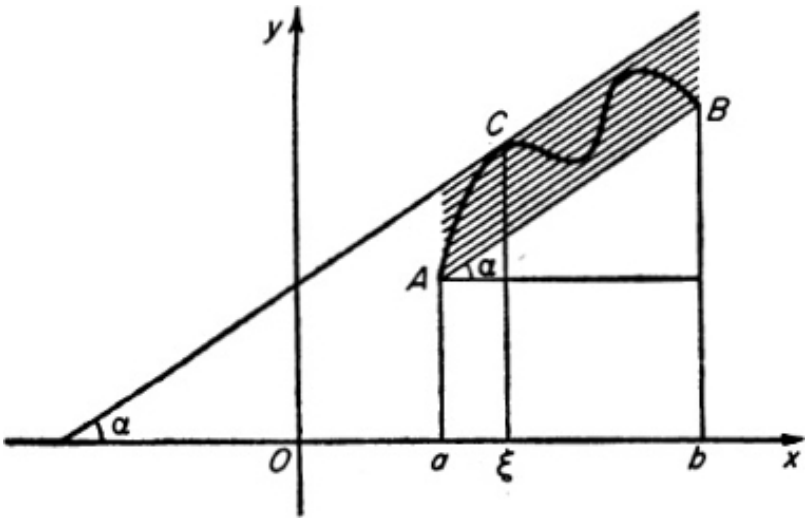


Figura 23.

$$\tan \alpha = \frac{f(b) - f(a)}{b - a}.$$

D'altra parte nel punto *C*

$$\tan \alpha = f'(\xi).$$

Questa equazione

$$\frac{f(b) - f(a)}{b - a} = f'(\xi)$$

è esattamente il teorema del valore medio. [19]

La formula (22) ha la particolarità che il punto ξ che vi compare ci è sconosciuto; sappiamo solo che si trova "da qualche parte nell'intervallo (*a*, *b*)." Ma nonostante questa indeterminatezza, la formula ha un grande significato teorico e fa parte della dimostrazione di molti teoremi in analisi. Molto grande è anche l'importanza pratica immediata di questa formula, poiché ci permette di stimare l'incremento di una funzione quando conosciamo i limiti entro i quali può variare la sua derivata. Per esempio,

$$|\sin b - \sin a| = |\cos \xi|\,(b - a) \leqslant b - a.$$

Qui *a*, *b* e ξ sono angoli, espressi in radianti; ξ è un valore compreso tra *a* e *b*; ξ stesso è sconosciuto, ma sappiamo che $|\cos \xi| \leqslant 1$.

Dalla formula (22) è chiaro che una funzione la cui derivata è ovunque uguale a zero deve essere una costante; in nessuna parte dell'intervallo può ricevere un incremento diverso da zero. Analogamente, il lettore dimostrerà facilmente che una funzione la cui derivata è ovunque positiva deve ovunque aumentare, e se la sua derivata è negativa la funzione deve diminuire. Diamo qui senza dimostrazione una delle tante generalizzazioni del teorema del valore medio.

Per funzioni arbitrarie $\phi(x)$ e $\psi(x)$ derivabili nell'intervallo $[a, b]$, a condizione solo che $\psi'(x) \neq 0$ in (a, b), vale la seguente equazione[20]

$$\frac{\phi(b) - \phi(a)}{\psi(b) - \psi(a)} = \frac{\phi'(\xi)}{\psi'(\xi)}, \qquad (23)$$

dove ξ è un punto nell'intervallo (a, b).[21]

Da questo teorema possiamo derivare un metodo generale per calcolare i limiti di un'espressione simile

$$\lim_{x \to 0} \frac{\phi(x)}{\psi(x)}, \qquad (24)$$

se $\phi(0) = \psi(0) = 0$. Dalla formula(23) abbiamo

$$\frac{\phi(x)}{\psi(x)} = \frac{\phi(x) - \phi(0)}{\psi(x) - \psi(0)} = \frac{\phi'(\xi)}{\psi'(\xi)},$$

dove ξ è compreso tra 0 e x, e quindi $\xi \to 0$ insieme a x. Questo ci permette di calcolare il limite

$$\lim_{x \to 0} \frac{\phi'(x)}{\psi'(x)},$$

invece del limite (24), che in molti casi è molto più semplice.[22]

Esempio. Troviamo lo $\lim_{x \to 0} \frac{x - \sin x}{x^3}$. Utilizzando la regola tre volte, abbiamo successivamente

$$\lim_{x \to 0} \frac{x - \sin x}{x^3} = \lim_{x \to 0} \frac{1 - \cos x}{3x^2} = \lim_{x \to 0} \frac{\sin x}{6x} = \lim_{x \to 0} \frac{\cos x}{6} = \frac{1}{6}.$$

§ 9. Formula di Taylor

La funzione

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

dove i coefficienti a_k sono costanti, si dice polinomio di grado n. In particolare $y = ax + b$ è un polinomio di primo grado e $y = ax^2 + bx + c$ è un polinomio di secondo grado. I polinomi possono essere considerati la più semplice di tutte le funzioni. Per calcolare il loro valore per un dato x, sono necessarie solo le operazioni di addizione, sottrazione e moltiplicazione; non è necessaria nemmeno la divisione. I polinomi sono continui per tutti gli x e hanno derivate di ordine arbitrario. Inoltre, la derivata di un polinomio è ancora un polinomio, di grado inferiore a uno, e le derivate di ordine $n + 1$ e superiori di un polinomio di grado n sono uguali a zero.

Se ai polinomi aggiungiamo funzioni della forma

$$y = \frac{a_0 + a_1x + \cdots + a_nx^n}{b_0 + b_1x + \cdots + b_mx^m},$$

per il cui calcolo occorre anche la divisione, e anche le funzioni $\sqrt{x}$ e $\sqrt[3]{x}$ e, infine, combinazioni aritmetiche di queste funzioni, si ottengono sostanzialmente tutte le funzioni i cui valori possono essere calcolati con metodi appresi nella scuola secondaria.

Mentre eravamo ancora a scuola, ci siamo fatti un'idea di una serie di altre funzioni, come

$$\sqrt[5]{x}, \log x, \sin x, \arctan x, \cdots .$$

Ma anche se abbiamo conosciuto le proprietà più importanti di queste funzioni, non abbiamo trovato risposta nella matematica elementare alla domanda: come possiamo calcolarle? Che tipo di operazioni, ad esempio, è necessario eseguire su x per ottenere $\log x$ o $\sin x$? La risposta a questa domanda è data dai metodi che sono stati elaborati nell'analisi. Esaminiamo uno di questi metodi.

La formula di Taylor. Su un intervallo contenente il punto a, sia data una funzione $f(x)$ con derivate di ogni ordine. Il polinomio di primo grado

$$p_1(x) = f(a) + f'(a)\,(x - a)$$

ha lo stesso valore di $f(x)$ nel punto $x = a$ ed inoltre, come è facilmente verificabile, ha la stessa derivata di $f(x)$ in questo punto. Il suo grafico è una linea retta, tangente al grafico di $f(x)$ fino al punto a. È possibile scegliere un polinomio di secondo grado, vale a dire

$$p_2(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2}(x - a)^2,$$

che nel punto $x = a$ ha con $f(x)$ un valore comune e una derivata prima e seconda comuni. Il suo grafico nel punto a seguirà ancora più da vicino quello di $f(x)$. È naturale aspettarsi che se costruiamo un polinomio che in $x = a$ ha le stesse prime derivate n di $f(x)$ nello stesso punto, allora questo polinomio sarà un'approssimazione ancora migliore di $f(x)$ in punti x vicino a a. Otteniamo così la seguente uguaglianza approssimativa, che è la formula di Taylor

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n . \quad (25)$$

Il lato destro di questa formula è un polinomio di grado n in $(x - a)$. Per ogni x si può calcolare il valore di questo polinomio se conosciamo i valori di $f(a), f'(a), \cdots, f^{(n)}(a)$.

Per funzioni che hanno una derivata $(n + 1)$esima, il lato destro di questa formula, come è facile mostrare, differisce dal lato sinistro per una piccola quantità che si avvicina allo zero più rapidamente di $(x - a)^n$. Inoltre, è l'unico polinomio possibile di grado n che differisce da $f(x)$, per x vicino a a, per una quantità che si avvicina allo zero, poiché $x \to a$, più rapidamente di $(x - a)^n$. Se $f(x)$ stesso è un polinomio algebrico di grado n, allora l'uguaglianza approssimativa (25) diventa esatta.

Infine, e questo è particolarmente importante, possiamo dare una semplice espressione per la differenza tra la parte destra della formula (25) e il valore effettivo di $f(x)$. Per rendere esatta l'uguaglianza approssimata (25), dobbiamo aggiungere a destra un ulteriore termine, chiamato "termine residuo"

$$f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1} \quad (26)$$

Questo termine supplementare finale[23]

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

ha la particolarità che la derivata che vi figura deve essere calcolata non nel punto a ma in un punto ξ opportunamente scelto, sconosciuto ma situato nell'intervallo tra a e x.

La prova dell'uguaglianza (26) è piuttosto complessa ma nella sua essenza piuttosto semplice. Daremo qui una versione un po' artificiale della dimostrazione, che ha il pregio di essere concisa.

Per scoprire di quanto il lato sinistro nella formula approssimata (25) differisce da quello destro, consideriamo il rapporto tra la differenza tra i due lati in uguaglianza (25) e la quantità $-(x - a)^{n+1}$

$$\frac{f(x) - \left[f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n\right]}{-(x-a)^{n+1}}. \qquad (27)$$

Introduciamo anche la funzione

$$\phi(u) = f(u) + f'(u)(x-u) + \cdots + \frac{f^{(n)}(u)}{n!}(x-u)^n$$

di una variabile u, considerando x fissa (costante). Quindi il numeratore nella (27) rappresenterà l'aumento di questa funzione mentre passiamo da $u = a$ a $u = x$, e il denominatore sarà l'aumento nello stesso intervallo della funzione

$$\psi(u) = (x-u)^{n+1}.$$

Utilizziamo ora il teorema del valore medio generalizzato citato in precedenza

$$\frac{\phi(x) - \phi(a)}{\psi(x) - \psi(a)} = \frac{\phi'(\xi)}{\psi'(\xi)}.$$

Differenziando le funzioni $\phi(u)$ e $\psi(u)$ rispetto a u (ricordiamo che il valore di x è stato fissato) troviamo che

$$\frac{\phi'(\xi)}{\psi'(\xi)} = -\frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

L'uguaglianza di quest'ultima espressione con la quantità originaria (27) dà la formula di Taylor nella forma (26).

Nella forma (26) la formula di Taylor non solo fornisce un mezzo di calcolo approssimativo di $f(x)$ ma ci permette anche di stimare l'errore. Consideriamo il semplice esempio

$$y = \sin x.$$

I valori della funzione sin x e delle sue derivate di ordine arbitrario sono noti per $x = 0$. Utilizziamo questi valori per scrivere la formula di Taylor per sin x, scegliendo $a = 0$ e limitandoci al caso $n = 4$. Troviamo successivamente

$$\begin{aligned}
&f(x) = \sin x, && f'(x) = \cos x, && f''(x) = -\sin x,\\
&f'''(x) = -\cos x, && f^{IV}(x) = \sin x, && f^{V}(x) = \cos x;\\
&f(0) = 0 && f'(0) = 1, && f''(0) = 0,\\
&f'''(0) = -1. && f^{IV}(0) = 0, && f^{V}(\xi) = \cos \xi.
\end{aligned}$$

Perciò

$$\sin x = x - \frac{x^3}{6} + R_5, \quad \text{where} \quad R_5 = \frac{x^5}{120}\cos \xi.$$

Sebbene il valore esatto R_5 sia sconosciuto, possiamo comunque facilmente stimarlo dal fatto che $|\cos \xi| \leqslant 1$. Per tutti i valori di x compresi tra 0 e $\pi/4$ abbiamo

$$|R_5| = \left|\frac{x^5}{120}\cos \xi\right| < \frac{1}{120}\left(\frac{\pi}{4}\right)^5 < \frac{1}{400}.$$

Di conseguenza, sull'intervallo $[0, \pi/4]$ la funzione $\sin x$ può essere considerata, con precisione fino a $\frac{1}{400}$, come uguale al polinomio di terzo grado

$$\sin x = x - \frac{1}{6}x^3.$$

Se dovessimo prendere più termini nell'espansione di Taylor per il peccato x, otterremmo un polinomio di grado più alto che approssimerebbe ancora più da vicino il peccato x.

Le tabelle per le funzioni trigonometriche e altre funzioni vengono calcolate con metodi simili.

Le leggi della natura, di regola, possono essere espresse con buona approssimazione da funzioni che possono essere differenziate quante volte si vuole e che a loro volta possono essere approssimate da polinomi, il grado del polinomio essendo determinato dalla precisione desiderata.

La serie di Taylor. Se nella formula (25) prendiamo un numero sempre maggiore di termini, allora la differenza tra il lato destro e $f(x)$, espressa dal termine rimanente $R_{n+1}(x)$, può tendere a zero. Naturalmente questo non si verificherà sempre: né per tutte le funzioni né per tutti i valori di x. Ma esiste un'ampia classe di funzioni (le cosiddette funzioni analitiche) per le quali il termine residuo $R_{n+1}(x)$ si avvicina effettivamente a zero come $n \to \infty$, almeno per tutti i valori di x entro un certo intervallo attorno al punto a. Per queste funzioni la formula di Taylor ci consente di calcolare $f(x)$ con qualsiasi grado di precisione desiderato. Esaminiamo tali funzioni più da vicino.

Se $R_{n+1}(x) + 0$ come $n \to \infty$, allora dalla (26) segue che

$$f(x) = \lim_{n\to\infty}\left[f(a) + f'(a)(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n\right]$$

In questo caso diciamo che $f(x)$ è stato espanso in una serie infinita convergente

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots,$$

in potenze crescenti di $(x - a)$. Questa serie è chiamata serie di Taylor e si dice che $f(x)$ sia la somma della serie. Consideriamo alcuni esempi

1.

$$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \cdots$$

(valid for $|x| < 1$ and for arbitrary real n).

Condizione: |x| < 1, con n reale arbitrario.

2.

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \qquad \text{(valid for all } x\text{)}.$$

Valida per ogni x.

3.

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \qquad \text{(valid for all } x\text{)}.$$

Valida per ogni x.

4.

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \qquad \text{(valid for all } x\text{)}.$$

Valida per ogni x.

5.

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \qquad (\text{valid for } |x| \leqslant 1).$$

Condizione: |x| ≤ 1.

Il primo di questi esempi è il famoso teorema binomiale di Newton, che fu ottenuto da Newton per tutti gli *n* ma completamente dimostrato a suo tempo solo per *n* intero. Questo esempio è servito da modello per stabilire la formula generale di Taylor. Le ultime due formule ci permettono, per $x = 1$, di calcolare con approssimazione arbitrariamente buona i numeri e e π.

La formula di Taylor, che apre la strada alla maggior parte dei calcoli dell'analisi applicata, è estremamente importante dal punto di vista pratico.

Molte delle leggi della natura, dei processi fisici e chimici, del movimento dei corpi e simili, sono espressi con grande accuratezza da funzioni che possono essere sviluppate in una serie di Taylor. La teoria di tali funzioni può essere formulata in modo più chiaro e completo se le consideriamo come funzioni di una variabile complessa (vedi Capitolo IX).

L'idea di approssimare una funzione mediante polinomi o di rappresentarla come somma di un numero infinito di funzioni più semplici ha subito profondi sviluppi in analisi, dove costituisce ora un ramo indipendente, la teoria dell'approssimazione delle funzioni (vedi capitolo XII).

§ 10. Integrale

Dal Capitolo I e dal §1 del presente capitolo il lettore sa già che il concetto di integrale, e più in generale di calcolo integrale, ha avuto la sua origine storica nella necessità di risolvere problemi concreti, di cui un esempio caratteristico è il calcolo dell'area di una figura curvilinea. A queste domande è dedicata la presente sezione. In essa si discuterà anche del suddetto collegamento tra i problemi del calcolo differenziale e del calcolo integrale, che non fu del tutto chiarito fino al XVIII secolo.

Zona. Supponiamo che una curva sopra l'asse x formi il grafico della funzione $y = f(x)$. Cerchiamo di trovare l'area S del segmento delimitato dalla linea $y = f(x)$, dall'asse x e dalle rette passanti per i punti $x = a$ e $x = b$ parallelo all'asse y.

Per risolvere questo problema procediamo come segue. Dividiamo l'intervallo $[a, b]$ in parti n, non necessariamente uguali. Indicheremo la lunghezza della prima parte con Δx_1, della seconda con Δx_2, e così via fino alla parte finale Δx_n. In ogni segmento scegliamo i punti ξ_1, ξ_2, $\cdots$, ξ_n e impostiamo la somma

$$S_n = f(\xi_1)\,\Delta x_1 + f(\xi_1)\,\Delta x_2 + \cdots + f(\xi_n)\,\Delta x_n . \qquad (28)$$

La grandezza S_n è ovviamente pari alla somma delle aree dei rettangoli ombreggiati in figura 24.

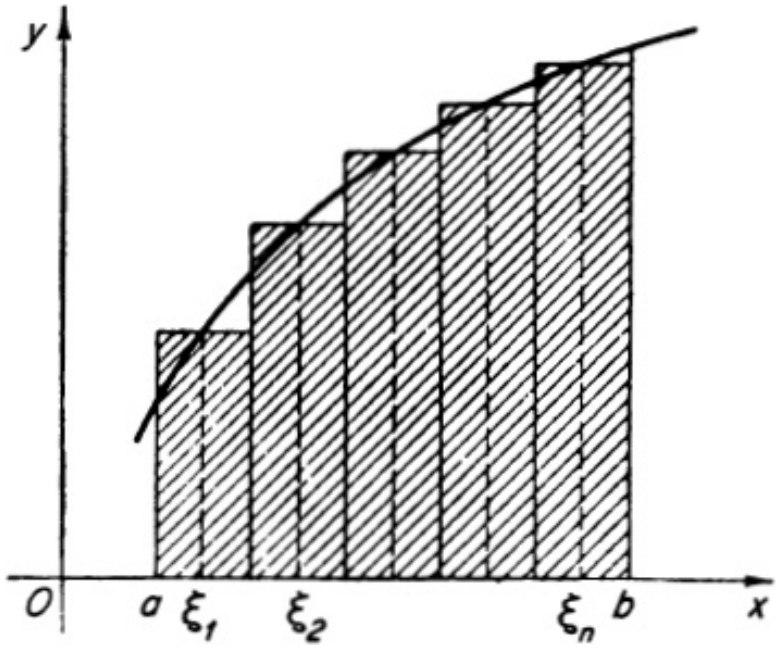


Figura 24.

Quanto più precisa sarà la suddivisione del segmento $[a, b]$, tanto più S_n si avvicinerà all'area S. Se eseguiamo una successione di tali costruzioni, dividendo l'intervallo $[a, b]$ in parti successivamente sempre più piccole, allora le somme S_n si avvicineranno a S.

La possibilità di dividere [*a*, *b*] in parti disuguali ci rende necessario definire cosa intendiamo per suddivisioni “successivamente più piccole”. Assumiamo non solo che *n* aumenti oltre ogni limite, ma anche che la lunghezza del massimo Δx_i nella suddivisione *n*-esima si avvicini allo zero. Così

$$S = \lim_{\max \Delta x_i \to 0} [f(\xi_1)\,\Delta x_1 + f(\xi_2)\,\Delta x_2 + \cdots + f(\xi_n)\,\Delta x_n]$$

$$= \lim_{\max \Delta x_i \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i\,. \qquad (29)$$

Il calcolo dell'area desiderata si è così ridotto alla ricerca del limite (29).

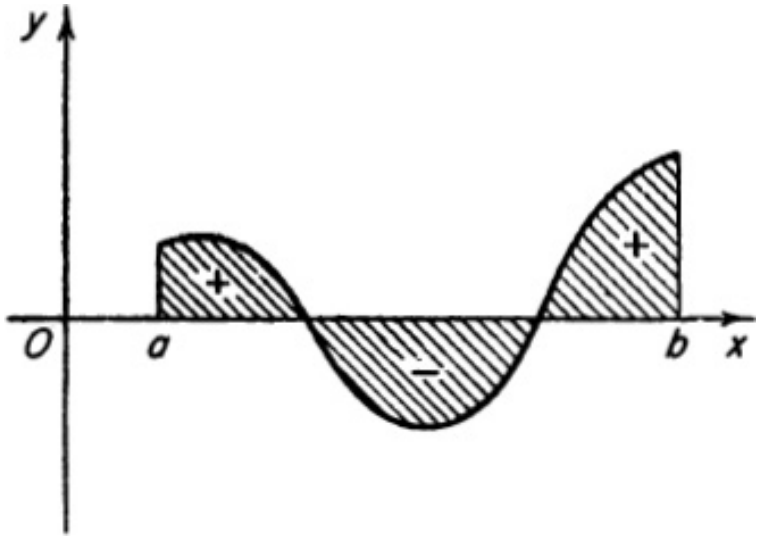


Figura 25.

Notiamo che quando abbiamo posto per la prima volta il problema, avevamo solo un'idea empirica di cosa intendiamo per area della nostra figura curvilinea, ma non avevamo una definizione precisa. Ma ora abbiamo ottenuto una definizione esatta del concetto di area: È il limite (29). Ora abbiamo non solo una nozione intuitiva di area ma anche una definizione matematica, in base alla quale possiamo calcolare numericamente l'area (confrontiamo le osservazioni alla fine del §3, riguardanti la velocità e la lunghezza di una circonferenza).

Abbiamo ipotizzato che $f(x) \geqslant 0$. Se $f(x)$ cambia segno, allora nella figura 25, il limite (29) ci darà la somma algebrica delle aree dei segmenti compresi tra la curva $y = f(x)$ e l'asse x, dove i segmenti sopra l'asse Gli assi x vengono presi con un segno più e quelli sottostanti con un segno meno.

Integrale definito. La necessità di calcolare il limite (29) sorge in molti altri problemi. Ad esempio, supponiamo che un punto si muova lungo una linea retta con velocità variabile $v = f(t)$. Come determiniamo la distanza s percorsa dal punto nel tempo da $t = a$ a $t = b$?

Supponiamo che la funzione $f(t)$ sia continua; cioè in piccoli intervalli di tempo la velocità cambia solo leggermente. Dividiamo l'intervallo [*a*, *b*] in parti *n*, di lunghezza $\Delta t_1, \Delta t_2, \cdots, \Delta t_n$. Per calcolare un valore approssimativo della distanza percorsa in ciascun intervallo Δt_i,

supponiamo che la velocità in questo periodo di tempo sia costante, uguale ovunque al suo valore effettivo in un punto intermedio ξ_1. L'intera distanza percorsa sarà quindi espressa approssimativamente dalla somma

$$s_n = \sum_{i=1}^{n} f(\xi_i)\,\Delta t_i\,,$$

e il valore esatto della distanza s percorsa nel tempo da a a b, costituirà il limite di tali somme per suddivisioni sempre più fini; cioè sarà il limite (29)

$$s = \lim_{\max \Delta t_i \to 0} \sum_{i=1}^{n} f(\xi_i)\,\Delta t_i\,.$$

Sarebbe facile fornire molti esempi di problemi pratici che portano al calcolo di tale limite. Ne discuteremo alcuni in seguito, ma per il momento gli esempi già forniti indicheranno sufficientemente l'importanza di questa idea. Il limite (29) è chiamato integrale definito della funzione $f(x)$ preso nell'intervallo $[a, b]$, ed è indicato con

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

L'espressione $f(x)dx$ è detta integrando, a e b sono i limiti dell'integrazione; a è il limite inferiore, b è il limite superiore.

La connessione tra calcolo differenziale e integrale. Come esempio di calcolo diretto di un integrale definito possiamo prendere l'esempio 2, §1. Possiamo ora dire che il problema qui considerato si riduce al calcolo dell'integrale definito

$$\int_0^h ax\,dx.$$

Un altro esempio è stato considerato nel §3, dove abbiamo risolto il problema di trovare l'area delimitata dalla parabola $y = x^2$. Qui il problema si riduce al calcolo dell'integrale

$$\int_0^1 x^2\,dx.$$

Siamo stati in grado di calcolare direttamente entrambi questi integrali, perché disponiamo di formule semplici per la somma dei primi numeri naturali n e per la somma dei loro quadrati. Ma per una funzione arbitraria $f(x)$, siamo lontani dal poter sommare la somma (28) (cioè esprimere il risultato in una formula semplice) se i punti ξ_i e gli incrementi Δx_i sono dati per soddisfare qualche problema particolare. Inoltre, anche quando tale somma è possibile, non esiste un metodo generale per effettuarla; nei vari casi devono essere utilizzati metodi diversi, ciascuno di carattere del tutto particolare.

Ci troviamo quindi di fronte al problema di trovare un metodo generale per il calcolo degli integrali defiqiti. Storicamente questa questione ha interessato i matematici per un lungo

periodo di tempo, poiché molti aspetti pratici erano coinvolti in un metodo generale per trovare l'area delle figure curvilinee, il volume dei corpi delimitati da una superficie curva, e così via.

Abbiamo già notato che Archimede era in grado di calcolare l'area di un segmento e di alcune altre figure. Il numero di problemi particolari che potevano essere risolti, riguardanti aree, volumi, centri di gravità dei solidi, e così via, aumentò gradualmente, ma i progressi nella ricerca di un metodo generale furono inizialmente estremamente lenti. Il metodo generale non poteva essere scoperto finché non fosse stato accumulato sufficiente materiale teorico e computazionale attraverso le esigenze della vita pratica. L'opera di raccolta e generalizzazione di questo materiale procedette molto gradualmente fino alla fine del Medioevo; e il suo successivo sviluppo energetico fu una conseguenza diretta della rapida crescita delle forze produttive dell'Europa risultante dalla disgregazione dei vecchi metodi di produzione (feudali) e dalla creazione di nuovi (capitalistici).

L'accumulazione dei fatti relativi agli integrali definiti procedette parallelamente alle corrispondenti indagini sui problemi legati alla derivata di una funzione. Il lettore sa già dal §1 che questo immenso lavoro preparatorio fu coronato con successo nel XVII secolo dall'opera di Newton e Leibniz. È in questo senso che Newton e Leibniz sono gli ideatori del calcolo differenziale e integrale.

Uno dei contributi fondamentali di Newton e Leibniz consiste nel fatto che essi hanno finalmente chiarito il profondo legame tra calcolo differenziale e calcolo integrale, il che ci fornisce, in particolare, un metodo generale per il calcolo degli integrali definiti per una classe estremamente ampia di funzioni.

Per spiegare questa connessione, prendiamo un esempio dalla meccanica.

Supponiamo che un punto materiale si muova lungo una linea retta con velocità $v = f(t)$, dove t è il tempo. Sappiamo già che la distanza a percorsa dal nostro punto nel tempo compreso tra $t = t_1$ e $t = t_2$ è data dall'integrale definito

$$\sigma = \int_{t_1}^{t_2} f(t)\, dt.$$

Supponiamo ora che la legge di moto del punto ci sia nota; cioè conosciamo la funzione $s = F(t)$ che esprime la dipendenza dal tempo t della distanza s calcolata a partire da un punto iniziale A sulla retta. La distanza σ percorsa nell'intervallo di tempo $[t_1, t_2]$ è ovviamente pari alla differenza

$$\sigma = F(t_2) - F(t_1).$$

In questo modo siamo condotti da considerazioni fisiche all'uguaglianza

$$\int_{t_1}^{t_2} f(t)\, dt = F(t_2) - F(t_1),$$

che esprime la connessione tra la legge del moto del nostro punto e la sua velocità.

Da un punto di vista matematico la funzione $F(t)$, come già sappiamo dal §5, può essere definita come una funzione la cui derivata per tutti i valori spesso nell'intervallo dato è uguale a $f(t)$, cioè

$$F'(t) = f(t).$$

Tale funzione è chiamata primitiva per $f(t)$.

Dobbiamo tenere presente che se la funzione $f(t)$ ha almeno una primitiva, allora insieme a questa ne avrà infinite altre; poiché se $F(t)$ è una primitiva per $f(t)$, allora anche $F(t) + C$, dove C è una costante arbitraria, è una primitiva. Inoltre in questo modo esauriamo l'intero insieme di primitive per $f(t)$, poiché se $F_1(t)$ e $F_2(t)$ sono primitive per la stessa funzione $f(t)$, allora la loro differenza $\phi(t) = F_1(t) - F_2(t)$ ha una derivata $\phi'(t)$ che è uguale a zero in ogni punto di un dato intervallo in modo che $\phi(t)$ sia una costante.[24]

Dal punto di vista fisico i diversi valori della costante C determinano leggi del moto che differiscono tra loro solo per il fatto che corrispondono a tutte le possibili scelte per il punto iniziale del moto.

Siamo così portati al risultato che per una classe estremamente ampia di funzioni $f(x)$, comprendente tutti i casi in cui la funzione $f(x)$ può essere considerata come la velocità di un punto all'istante x, si ha la seguente uguaglianza[25]

$$\int_a^b f(x)\,dx = F(b) - F(a), \qquad (30)$$

dove $F(x)$ è una primitiva arbitraria per $f(x)$.

Questa uguaglianza è la famosa formula di Newton e Leibniz, che riduce il problema del calcolo dell'integrale definito di una funzione alla ricerca di una primitiva per la funzione e costituisce in questo modo un collegamento tra il calcolo differenziale e il calcolo integrale.

Molti problemi particolari studiati dai più grandi matematici vengono risolti da questa formula: l'integrale definito della funzione $f(x)$ sull'intervallo $[a, b]$ è uguale alla differenza fra i valori di una qualsiasi primitiva all'estremo destro e all'estremo sinistro dell'intervallo.[26] Si usa scrivere la differenza (30) nella forma:

$$F(x)\Big|_a^b = F(b) - F(a).$$

Esempio 1. L'uguaglianza

$$\left(\frac{x^3}{3}\right)' = x^2$$

mostra che la funzione $x^3/3$ è una primitiva per la funzione x^2. Quindi, con la formula di Newton e Leibniz,

$$\int_0^a x^2\,dx = \frac{x^3}{3}\Big|_0^a = \frac{a^3}{3} - \frac{0}{3} = \frac{a^3}{3}.$$

Esempio 2. Siano c e c' due cariche elettriche, allineate a distanza r l'una dall'altra. L'attrazione F tra loro è diretta lungo questa linea retta ed è uguale a

$$F = \frac{a}{r^2}$$

(a = kcc′, dove k è una costante). Il lavoro W compiuto da questa forza, quando la carica c rimane fissa ma c' si muove lungo l'intervallo $[R_1, R_2]$, può essere calcolato dividendo l'intervallo $[R_1, R_2]$ nelle parti Δr_i. Su ciascuna di queste parti possiamo considerare la forza approssimativamente costante, per cui il lavoro compiuto su ciascuna parte è pari a $a/r_i^2\,\Delta r_i$. Facendo le parti sempre più piccole, vediamo che il lavoro W è uguale all'integrale

$$W = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{a}{r_i^2}\,\Delta r_i = \int_{R_1}^{R_2} \frac{a}{r^2}\,dr.$$

Il valore di questo integrale può essere calcolato immediatamente, se lo ricordiamo

$$\frac{a}{r^2} = \left(-\frac{a}{r}\right)',$$

affinché

$$W = -\frac{a}{r}\Big|_{R_1}^{R_2} = a\left(\frac{1}{R_1} - \frac{1}{R_2}\right).$$

In particolare, il lavoro compiuto da una forza F quando la carica c', inizialmente a distanza R_1, da c, si sposta all'infinito, è pari a

$$W = \lim_{R_2\to\infty} a\left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \frac{a}{R_1}.$$

Dalle argomentazioni sopra esposte per la formula di Newton e Leibniz risulta chiaro che questa formula dà espressione matematica ad un legame reale esistente nel mondo oggettivo. È un esempio bello e importante di come la matematica dà espressione a leggi oggettive. Dobbiamo notare che nelle sue ricerche matematiche Newton adottò sempre un punto di vista fisico. Il suo lavoro sui fondamenti del calcolo differenziale e integrale non può essere separato dal suo lavoro sui fondamenti della meccanica.

I concetti dell'analisi matematica, come la derivata o l'integrale, come si presentavano a Newton e ai suoi contemporanei, non si erano ancora “separati” completamente dalle loro origini fisiche e geometriche, come velocità e area. In effetti, erano per metà di carattere matematico e per metà fisico. Le condizioni esistenti a quel tempo non erano ancora adatte per produrre una definizione puramente matematica di questi concetti. Di conseguenza, il ricercatore potrebbe gestirli correttamente in situazioni complicate solo se rimanesse in

stretto contatto con gli aspetti pratici del suo problema anche durante le fasi intermedie (matematiche) della sua argomentazione.

Da questo punto di vista il lavoro creativo di Newton fu diverso da quello di Leibniz.[27] Newton fu guidato in tutte le fasi da un modo fisico di considerare il problema. Ma le ricerche di Leibniz non hanno un legame così immediato con la fisica, fatto che, in assenza di definizioni matematiche chiare, lo portava talvolta a conclusioni errate. D'altra parte, la caratteristica più caratteristica dell'attività creativa di Leibniz era la sua ricerca della generalità, i suoi sforzi per trovare i metodi più generali per i problemi dell'analisi matematica.

Il merito più grande di Leibniz fu la creazione di un simbolismo matematico che esprimesse l'essenza della questione. Le notazioni per concetti fondamentali dell'analisi matematica come il differenziale dx, il secondo differenziale d_2x, l'integrale $\int y\,dx$ e la derivata d/dx sono state proposte da Leibniz. Il fatto che queste notazioni siano ancora utilizzate dimostra quanto siano state scelte bene.

Un vantaggio di un simbolismo ben scelto è che rende le nostre dimostrazioni e i nostri calcoli più brevi e più facili; inoltre, a volte ci protegge da conclusioni errate. Leibniz, che ne era ben consapevole, prestò in tutta la sua opera una particolare attenzione alla scelta della notazione.

L'evoluzione dei concetti di analisi matematica (derivata, integrale, ecc.) è continuata, ovviamente, dopo Newton e Leibniz e continua ancora ai nostri giorni; ma c'è una fase in questa evoluzione che merita di essere menzionata in particolare. Ha avuto luogo all'inizio del secolo scorso ed è legato in particolare all'opera di Cauchy.

Cauchy diede una definizione formale chiara del concetto di limite e la usò come base per le sue definizioni di continuità, derivata, differenziale e integrale. Queste definizioni sono state introdotte nei punti corrispondenti del presente capitolo. Sono ampiamente utilizzate nell'analisi odierna.

La grande importanza di queste conquiste sta nel fatto che ora è possibile operare in modo puramente formale non solo nell’aritmetica, nell’algebra e nella geometria elementare, ma anche in questo nuovo e molto vasto ramo della matematica, nell’analisi matematica, e ottenere così risultati corretti.

Per quanto riguarda l'applicazione pratica dei risultati dell'analisi matematica, ora si può dire: se i dati originali vengono verificati nel mondo reale, allora anche i risultati delle nostre argomentazioni matematiche saranno verificati lì. Se siamo adeguatamente certi dell'accuratezza dei dati originali, non è necessario effettuare un controllo pratico della correttezza dei risultati matematici; è sufficiente verificare solo la correttezza delle argomentazioni formali.

Questa affermazione richiede naturalmente la seguente limitazione. Nelle argomentazioni matematiche i dati originali, che prendiamo dal mondo reale, sono veri solo fino a una certa precisione. Ciò significa che ad ogni passo del nostro ragionamento matematico i risultati

ottenuti conterranno alcuni errori, che potrebbero accumularsi man mano che aumenta il numero di passi del ragionamento.[28]

Ritornando ora all'integrale definito, consideriamo una questione di fondamentale importanza. Per quali funzioni $f(x)$, definite sull'intervallo $[a, b]$, è possibile garantire l'esistenza dell'integrale definito $\int_a^b f(x)\,dx$, cioè un numero al quale la somma $\Sigma_1^n f(\xi_i)\,\Delta x_i$, tende come limite massimo $\Delta x_i \to 0$? Bisogna tenere presente che questo numero deve essere lo stesso per tutte le suddivisioni dell'intervallo $[a, b]$ e per tutte le scelte dei punti ξ_i.

Le funzioni per le quali esiste l'integrale definito, cioè il limite (29), si dicono integrabili sull'intervallo $[a, b]$. Le indagini effettuate nel secolo scorso mostrano che tutte le funzioni continue sono integrabili.

Ma ci sono anche funzioni discontinue che sono integrabili. Tra queste, ad esempio, ci sono quelle funzioni limitate e crescenti o decrescenti nell'intervallo $[a, b]$.

La funzione uguale a zero nei punti razionali in $[a, b]$ e uguale all'unità nei punti irrazionali, può servire come esempio di funzione non integrabile, poiché per una suddivisione arbitraria la somma integrale s_n, sarà uguale a zero o all'unità, a seconda che si scelgano i punti ξ_i, come numeri razionali o irrazionale.

Notiamo che in molti casi la formula di Newton e Leibniz fornisce una risposta alla questione pratica del calcolo di un integrale definito. Ma qui sorge il problema di trovare una primitiva per una data funzione; cioè di trovare una funzione che abbia come derivata la funzione data. Procediamo ora a discutere questo problema. Notiamo tra l'altro che il problema di trovare una primitiva ha grande importanza anche in altri rami della matematica, in particolare nella soluzione delle equazioni differenziali.

§ 11. Integrali indefiniti; tecniche di integrazione

Una primitiva arbitraria di una data funzione $f(x)$ è solitamente chiamata integrale indefinito di $f(x)$ ed è scritta nella forma

$$\int f(x)\,dx = F(x) + C, \qquad (31)$$

In questo modo, se $F(x)$ è una primitiva completamente determinata di $f(x)$, allora l'integrale indefinito di $f(x)$ è dato da

$$\int f(x)\,dx = F(x) + C, \qquad (31)$$

dove C è una costante arbitraria.

Notiamo inoltre che se la funzione $f(x)$ è data sull'intervallo $[a, b]$ e, se $F(x)$ è una primitiva per $f(x)$ e x è un punto nell'intervallo $[a, b]$, quindi con la formula di Newton e Leibniz possiamo scrivere

$$F(x) = F(a) + \int_a^x f(t)\,dt.$$

Qui l'integrale a destra differisce dalla primitiva $F(x)$ solo per la costante $F(a)$. In tal caso questo integrale, se lo consideriamo in funzione del suo estremo superiore x (per la variabile x), è una primitiva completamente determinata di $f(x)$. Di conseguenza, un integrale indefinito di $f(x)$ può anche essere scritto come segue:

$$\int f(x)\,dx = \int_a^x f(t)\,dt + C,$$

dove C è una costante arbitraria.

Costruiamo una tabella fondamentale degli integrali indefiniti, che si ottiene direttamente dalla corrispondente tabella delle derivate (vedi §6):

$$\left.\begin{aligned}
&\int x^a\,dx = \frac{x^{a+1}}{a+1} + C (a \neq -1),\\
&\int \frac{dx}{x} = \ln|x| + C,^*\\
&\int a^x\,dx = \frac{a^x}{\ln a} + C,\\
&\int e^x\,dx = e^x + C,\\
&\int \sin x\,dx = -\cos x + C,\\
&\int \cos x\,dx = \sin x + C,\\
&\int \sec^2 x\,dx = \tan x + C,\\
&\int \frac{dx}{\sqrt{1-x^2}} = \operatorname{arc\,sin} x + C\\
&\qquad\qquad = -\operatorname{arc\,cos} x + C_1 \left(C_1 - C = \frac{\pi}{2}\right),\\
&\int \frac{dx}{1+x^2} = \operatorname{arc\,tan} x + C.
\end{aligned}\right\} \quad (32)$$

Nota alla formula dell'integrale di $1/x$: [28a]

Le proprietà generali degli integrali indefiniti possono essere dedotte anche dalle corrispondenti proprietà delle derivate. Ad esempio, dalla regola per la derivazione di una somma si ottiene la formula

$$\int [f(x) \pm \phi(x)]\,dx = \int f(x)\,dx \pm \int \phi(x)\,dx + C,$$

e dalla corrispondente regola che esprime il fatto che un fattore costante k può essere preso fuori dal segno di derivazione si ottiene

$$\int kf(x)\,dx = k \int f(x)\,dx + C.$$

Per esempio,

$$\int \left(3x^2 + 2x - \frac{3}{\sqrt{x}} + \frac{4}{x} - 1\right) dx$$
$$= 3\frac{x^3}{3} + \frac{2x^2}{2} - 3\frac{x^{-1/2+1}}{-\frac{1}{2}+1} + 4\ln|x| - x + C.$$

Esistono diversi metodi per calcolare gli integrali indefiniti. Consideriamone uno, vale a dire il metodo di sostituzione o cambiamento di variabile, che si basa sulla seguente uguaglianza

$$\int f(x)\,dx = \int f[\phi(t)]\,\phi'(t)\,dt + C, \qquad (33)$$

dove $x = \phi(t)$ è una funzione derivabile. La relazione (33) è da intendersi nel senso che se nella funzione

$$F(x) = \int f(x)\,dx,$$

a sinistra dell'uguaglianza (33), poniamo $x = \phi(t)$, otteniamo così una funzione $F[\phi(t)]$ la cui derivata rispetto a t è uguale all'espressione sotto il segno di integrazione a destra dell'uguaglianza (33). Questo fatto segue immediatamente dal teorema sulla derivata di una funzione composta.

Diamo alcuni esempi di questo metodo di sostituzione

$$\int e^{kx}\,dx = \int e^t \frac{1}{k}\,dt = \frac{1}{k}\int e^t\,dt = \frac{1}{k}e^t + C = \frac{e^{kx}}{k} + C$$

(sostituzione $kx = t$, da cui $k\,dx = dt$).

$$\int \frac{d}{\sqrt{a^2 - x^2}} = -\int dt = -t + C = -\sqrt{a^2 - x^2} + C$$

(sostituzione $\sqrt{a^2 - x^2}$, da cui $-\frac{x\,dx}{\sqrt{a^2 - x^2}}$).

$$\begin{aligned}\int \sqrt{a^2 - x^2}\,dx &= \int \sqrt{a^2 - a^2\sin^2 u}\; a\cos u\,du = a^2\int \cos^2 u\,du \\ &= a^2\int \frac{1 + \cos 2u}{2}\,du = \frac{a^2}{2}\left(u + \frac{\sin 2u}{2}\right) + C \\ &= \frac{a^2}{2}(u + \sin u\cos u) + C \\ &= \frac{a^2}{2}\left(\arcsin\frac{x}{a} + \frac{x}{a^2}\sqrt{a^2 - x^2}\right) + C\end{aligned}$$

(sostituzione di $x = a\sin u$).

Come si vede da questi esempi, il metodo di sostituzione o cambiamento di variabili estende notevolmente la classe di funzioni elementari che possiamo integrare; cioè per le quali possiamo trovare primitive che sono esse stesse funzioni elementari. Ma va notato che, dal punto di vista del calcolo effettivo del risultato, siamo in una posizione molto peggiore, in generale, per quanto riguarda l'integrazione che per la derivazione.

Dal §6 sappiamo che la derivata di una funzione elementare arbitraria è essa stessa una funzione elementare, che possiamo effettivamente calcolare facendo uso delle regole di derivazione. Ma l'affermazione contraria è in generale falsa, poiché esistono funzioni elementari i cui integrali indefiniti non sono funzioni elementari. Esempi sono e^{-x^2}, $1/(\ln x)$, $(\sin x)/x$ e così via. Per ottenere gli integrali di queste funzioni dobbiamo utilizzare metodi approssimativi e introdurre anche nuove funzioni non riducibili a elementari. Non possiamo soffermarci ulteriormente qui su questa questione ma dobbiamo semplicemente notare che anche nella matematica elementare è possibile trovare molti esempi in cui un'operazione diretta può essere effettuata su una certa classe di numeri, mentre l'operazione inversa non

può essere effettuata sulla stessa classe; quindi, un quadrato di un numero razionale arbitrario è ancora una volta un numero razionale, ma la radice quadrata di un numero razionale non è affatto sempre razionale. Analogamente, la derivazione delle funzioni elementari produce una funzione anch'essa elementare, ma l'integrazione può portarci fuori dalla classe delle funzioni elementari.

Alcuni degli integrali che non possono essere espressi in termini di funzioni elementari hanno grande importanza in matematica e nelle sue applicazioni. Un esempio è

$$\int_0^x e^{-t^2}\,dt,$$

che gioca un ruolo molto importante nella teoria della probabilità (vedi Capitolo XI). Altri esempi sono gli integrali

$$\int_0^\phi \frac{d\theta}{\sqrt{1-k^2\sin^2\theta}} \quad \text{and} \quad \int_0^\phi \sqrt{1-k^2\sin^2\theta}\,d\theta\ (k^2<1),$$

che sono detti integrali ellittici rispettivamente di prima e di seconda specie. Al calcolo di questi integrali siamo condotti da un gran numero di problemi di fisica (vedi Capitolo V, §1, esempio 3). Le tabelle dettagliate di questi integrali per vari valori degli argomenti x e ϕ sono state calcolate con metodi approssimativi ma con grande precisione.

Bisogna sottolineare che la prova stessa del fatto che una data funzione elementare non può essere integrata in termini di funzioni elementari è in ogni caso piuttosto difficile. Tali domande hanno occupato l'attenzione di eminenti matematici nel secolo scorso e hanno svolto un ruolo importante nello sviluppo dell'analisi. Risultati fondamentali li ottenne qui Čebyšev, che diede una risposta completa alla questione di esprimere in termini di funzioni elementari gli integrali della forma

$$\int x^m(a+bx^s)^p\,dx,$$

dove m, s e p sono numeri razionali. Fino a lui erano note tre relazioni, ottenute da Newton, per gli esponenti m, s e p, che implicavano l'integrabilità di questo integrale in termini di funzioni elementari. Čebyšev ha dimostrato che in tutti gli altri casi l'integrale non può essere espresso in termini di funzioni elementari.

Introduciamo qui un altro metodo di integrazione, vale a dire l'integrazione per parti. Si basa sulla formula che già conosciamo

$$(uv)' = uv' + u'v,$$

per la derivata del prodotto delle funzioni u e v. Questa formula può anche essere scritta

$$uv' = (uv)' - u'v.$$

Integriamo ora i lati sinistro e destro, tenendo presente questo

$$\int (uv)' \, dx = uv + C.$$

Ora finalmente otteniamo l'uguaglianza

$$\int uv' \, dx = uv - \int u'v \, dx,$$

detta anche formula di integrazione per parti. Non abbiamo scritto la costante C poiché possiamo considerarla inclusa in uno degli integrali indefiniti presenti in questa equazione.

Introduciamo alcune applicazioni di questa formula. Supponiamo di dover calcolare $\int xe^x \, dx$. Qui prenderemo $u = x$ e $v' = e^x$, e quindi $u' = 1$, $v = e^x$, e di conseguenza

$$\int xe^x \, dx = xe^x - \int 1 \cdot e^x \, dx = xe^x - e^x + C.$$

Nell'integrale $\int \ln x \, dx$ conviene prendere $u = \ln x$, $v' = 1$, per cui $u' = 1/x$, $v = x$ e

$$\int \ln x \, dx = x \ln x - \int dx = x \ln x - x + C.$$

Nel seguente esempio caratteristico è necessario integrare due volte per parti e poi ricavare l'integrale desiderato dalle equazioni così ottenute:

$$\begin{aligned} \int e^x \sin x \, dx &= e^x \sin x - \int e^x \cos x \, dx \\ &= e^x \sin x - e^x \cos x - \int e^x \sin x \, dx, \end{aligned}$$

da cui

$$\int e^x \sin x \, dx = \frac{e^x}{2} (\sin x - \cos x) + C.$$

Terminiamo qui questa sezione; da esso il lettore avrà ricavato solo un'idea superficiale della teoria dell'integrazione. Non abbiamo prestato attenzione a molti metodi diversi in questa teoria. In particolare non abbiamo toccato qui l'interessantissima questione dell'integrazione delle frazioni razionali, teoria alla quale un importante contributo venne dato dal noto matematico e meccanico Ostrogradskij.

§ 12. Funzioni di più variabili

Finora abbiamo parlato solo di funzioni di una variabile, ma in pratica spesso è necessario occuparsi anche di funzioni dipendenti da due, tre o in generale molte variabili. Ad esempio, l'area di un rettangolo è una funzione

$$S = xy$$

della sua base x e della sua altezza y. Il volume di un parallelepipedo rettangolare è una funzione

$$v = xyz$$

delle sue tre dimensioni. La distanza tra due punti A e B è una funzione

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

delle sei coordinate di questi punti. La formula ben nota

$$pv = RT$$

esprime la dipendenza del volume v di una determinata quantità di gas dalla pressione p e dalla temperatura assoluta T.

Le funzioni di più variabili, come le funzioni di una variabile, sono in molti casi definite solo su una certa regione di valori delle variabili stesse. Ad esempio, la funzione

$$u = \ln(1 - x^2 - y^2 - z^2) \qquad (34)$$

è definito solo per i valori di x, y e z che soddisfano la condizione

$$x^2 + y^2 + z^2 < 1. \qquad (35)$$

(Per gli altri x, y, z i suoi valori non sono numeri reali.) L'insieme dei punti dello spazio le cui coordinate soddisfano la disuguaglianza (35) riempie ovviamente una sfera di raggio unitario con il centro nell'origine delle coordinate. I punti sul confine non sono compresi in questa sfera; la superficie della sfera è stata per così dire “staccata via”. Si dice che tale sfera sia aperta. La funzione (34) è definita solo per tali insiemi di tre numeri (x, y, z) come sono le coordinate dei punti nella sfera aperta G. Si è soliti enunciare questo fatto in modo conciso dicendo che la funzione (34) è definita sulla sfera G.

Facciamo un altro esempio. La temperatura di un corpo riscaldato in modo non uniforme V è funzione delle coordinate x, y, z dei punti del corpo. Questa funzione non è definita per tutti gli insiemi di tre numeri x, y, z ma solo per gli insiemi che sono coordinate di punti del corpo V.

Infine, come terzo esempio, consideriamo la funzione

$$u = \phi(x) + \phi(y) + \phi(z),$$

dove ϕ è una funzione di una variabile definita nell'intervallo [0, 1]. Ovviamente la funzione u è definita solo per insiemi di tre numeri (x, y, z) che sono coordinate di punti nel cubo:

$$0 \leqslant x \leqslant 1,\ 0 \leqslant y \leqslant 1, 0 \leqslant z \leqslant 1.$$

Diamo ora una definizione formale di funzione di tre variabili. Supponiamo che ci venga dato un insieme E di triple di numeri (x, y, z) (punti dello spazio). Se a ciascuna di queste triple di numeri (punti) di E corrisponde un numero definito u secondo qualche legge, allora u si dice che sia una funzione di x, y, z (del punto), definita sull'insieme delle triple di numeri (sui punti) E, fatto che così è scritto:

$$u = F(x, y, z).$$

Al posto di F possiamo scrivere anche altre lettere: f, ϕ, ψ.

In pratica l'insieme E sarà solitamente un insieme di punti, che riempiono un corpo o una superficie geometrica: sfera, cubo, anello e così via, e quindi diciamo semplicemente che la funzione è definita su questo corpo o superficie. Le funzioni di due, quattro e così via variabili sono definite in modo analogo.

Definizione implicita di una funzione. Notiamo che le funzioni di due variabili possono servire, in determinate circostanze, come mezzo utile per la definizione di funzioni di una variabile. Data una funzione $F(x, y)$ di due variabili impostiamo l'equazione

$$F(x, y) = 0. \qquad (36)$$

In generale, questa equazione definirà un certo insieme di punti (x, y) della superficie su cui la nostra funzione è uguale a zero. Tali insiemi di punti rappresentano solitamente curve che possono essere considerate come i grafici di una o più funzioni a un valore $y = \phi(x)$ o $x = \psi(y)$ di una variabile. In tal caso si dice che queste funzioni ad un valore siano definite implicitamente dall'equazione (36). Ad esempio, l'equazione

$$x^2 + y^2 - r^2 = 0$$

fornisce una definizione implicita di due funzioni di una variabile

$$y = +\sqrt{r^2 - x^2} \quad \text{and} \quad y = -\sqrt{r^2 - x^2}.$$

Ma è necessario tenere presente che un'equazione della forma (36) potrebbe non riuscire a definire alcuna funzione. Ad esempio, l'equazione

$$x^2 + y^2 + 1 = 0$$

ovviamente non definisce alcuna funzione reale, poiché nessuna coppia di numeri reali la soddisfa.

Rappresentazione geometrica. Le funzioni di due variabili possono sempre essere visualizzate come superfici mediante un sistema di coordinate spaziali. Quindi la funzione

$$z = f(x, y) \qquad (37)$$

è rappresentato in un sistema di coordinate rettangolari tridimensionale da una superficie, che è il luogo geometrico dei punti M le cui coordinate x, y, z soddisfano l'equazione (37) (figura 26).

Esiste un altro metodo, estremamente utile, per rappresentare la funzione (37), che ha trovato ampia applicazione nella pratica. Scegliamo una successione di numeri z_1, z_2, ⋯, e poi disegniamo su uno stesso piano Oxy le curve

$$z_1 = f(x, y), \; z_2 = f(x, y),$$

che sono le cosiddette linee di livello della funzione $f(x, y)$. Da un insieme di linee di livello, se corrispondono a valori di z sufficientemente vicini tra loro, è possibile formare un'immagine molto buona della variazione della funzione $f(x, y)$, così come dalle linee di livello di una carta topografica si può giudicare la variazione di altitudine della località.

La Figura 27 mostra una mappa delle linee di livello della funzione $z = x^2 + y^2$, il diagramma a destra indica come la funzione è costruita dalle sue linee di livello. Nel Capitolo III, figura 50, una mappa simile è disegnata per le linee di livello della funzione $z = xy$.

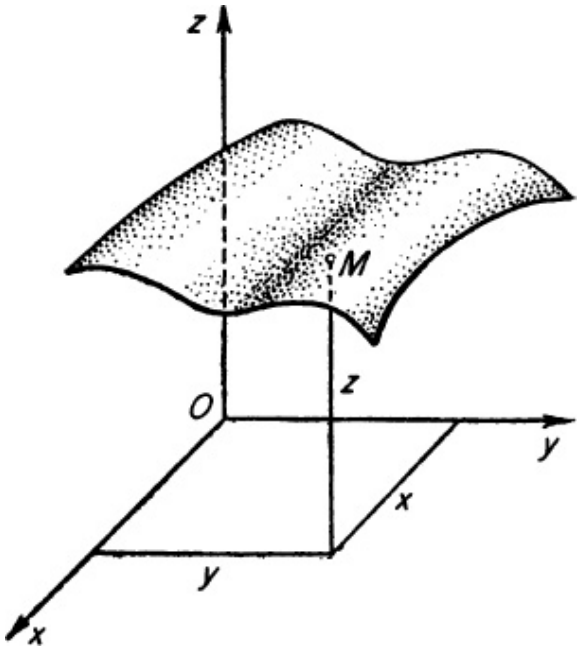


Figura 26.

Derivate parziali e differenziale. Facciamo alcune osservazioni sulla derivazione delle funzioni di più variabili. Prendiamo come esempio la funzione arbitraria

$$z = f(x, y)$$

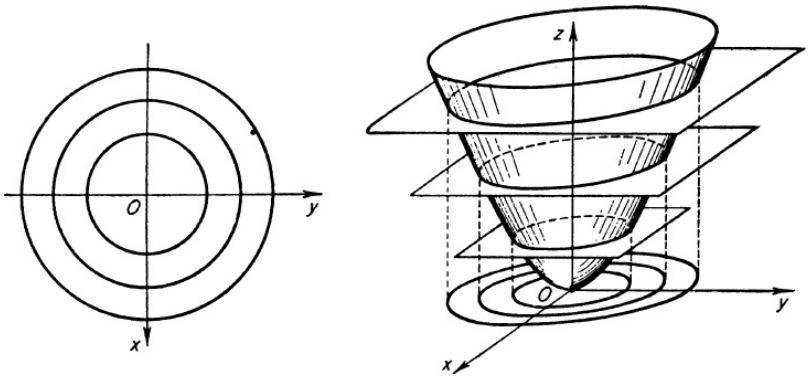


Figura 27.

di due variabili. Se fissiamo il valore di y, cioè se lo consideriamo non variabile, allora la nostra funzione di due variabili diventa una funzione dell'unica variabile x. La derivata di questa funzione rispetto a x, se esiste, è detta derivata parziale rispetto a x e si denota così:

$$\frac{\partial z}{\partial x}, \quad \text{or} \quad \frac{\partial f}{\partial x}, \quad \text{or} \quad f'_x(x, y).$$

L'ultima di queste tre notazioni indica chiaramente che la derivata parziale rispetto a x è in generale una funzione di x e y. La derivata parziale rispetto a y è definita in modo simile.

Dal punto di vista geometrico la funzione $f(x, y)$ rappresenta una superficie in un sistema di coordinate tridimensionale rettangolare. La corrispondente funzione di x per y fisso rappresenta una curva piana (figura 28) ottenuta dall'intersezione della superficie con un piano parallelo al piano Oxz e distante y da esso. La derivata parziale $\partial z/\partial x$ è ovviamente uguale alla tangente trigonometrica dell'angolo compreso tra le tangenti. alla curva nel punto (x, y) e alla direzione positiva dell'asse x.

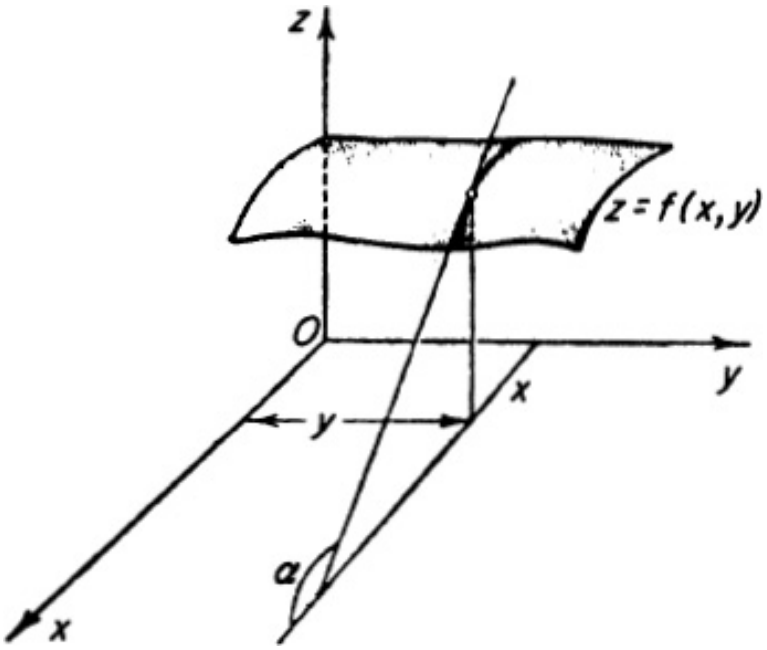


Figura 28.

Più in generale, se consideriamo una funzione $z = f(x_1, x_2, \ldots, x_n)$ delle variabili n $x_1, x_2, \ldots, x_n$, la derivata parziale $\partial z/\partial x_i$, è definita come la derivata di questa funzione rispetto a x_i, calcolata per valori fissi delle altre variabili:

$$x_1 , x_2 , \ldots, x_{i-1} , x_{i+1} , \ldots, x_n .$$

Possiamo dire che la derivata parziale di una funzione rispetto alla variabile x_i è la velocità di variazione di questa funzione nella direzione della variazione di x_i. Sarebbe anche possibile definire una derivata in una direzione assegnata arbitrariamente, non necessariamente coincidente con nessuno degli assi delle coordinate, ma non ci prenderemo il tempo per farlo.

Esempi

1.[1.] $z = \frac{x}{y}, \frac{\partial z}{\partial x} = \frac{1}{y}, \frac{\partial z}{\partial y} = -\frac{x}{y^2}$.

2.

$$u = \frac{1}{\sqrt{x^2 + y^2 + z^2}}$$

$$\frac{\partial u}{\partial x} = -\frac{1}{x^2 + y^2 + z^2} \cdot \frac{2x}{2\sqrt{x^2 + y^2 + z^2}} = -\frac{x}{(x^2 + y^2 + z^2)^{3/2}} .$$

Talvolta è necessario formare le derivate parziali di queste derivate parziali; cioè le cosiddette derivate parziali del secondo ordine. Per le funzioni di due variabili ce ne sono quattro

$$\frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x\, \partial y}, \frac{\partial^2 u}{\partial y\, \partial x}, \frac{\partial^2 u}{\partial y^2} .$$

Tuttavia, se queste derivate sono continue, allora non è difficile dimostrare che la seconda e la terza di queste quattro (le cosiddette derivate miste) coincidono:

$$\frac{\partial^2 u}{\partial x\, \partial y} = \frac{\partial^2 u}{\partial y\, \partial x} .$$

Ad esempio, nel caso della prima funzione considerata,

$$\frac{\partial^2 z}{\partial x^2} = 0, \frac{\partial^2 z}{\partial x\, \partial y} = -\frac{1}{y^2}, \frac{\partial^2 z}{\partial y\, \partial x} = -\frac{1}{y^2}, \frac{\partial^2 z}{\partial y^2} = \frac{2x}{y^3},$$

si vede che le due derivate miste coincidono.

Per le funzioni di più variabili, così come è stato fatto per le funzioni di una sola variabile, possiamo introdurre il concetto di differenziale.

Per chiarezza consideriamo una funzione

$$z = f(x, y)$$

di due variabili. Se ha derivate parziali continue, possiamo dimostrare che il suo incremento

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y),$$

corrispondente agli incrementi Δx e Δy dei suoi argomenti, può essere messo nella forma

$$\Delta z = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \alpha\sqrt{\Delta x^2 + \Delta y^2},$$

dove $\partial f/\partial x$ e $\partial f/\partial y$ sono le derivate parziali della funzione nel punto (x, y) e il modulo a dipende da Δx e Δy in tale modo che $\alpha \to 0$ come $\Delta x \to 0$ e $\Delta y \to 0$.

La somma dei primi due componenti

$$dz = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y$$

è linearmente dipendente[29] da Δx e Δy ed è chiamato differenziale della funzione. Il terzo addendo, per la presenza del fattore α, tendente a zero con Δx e Δy, è un infinitesimo di ordine superiore alla grandezza

$$\rho = \sqrt{\Delta x^2 + \Delta y^2},$$

descrivendo il cambiamento in x e y.

Diamo un'applicazione del concetto di differenziale. Il periodo di oscillazione di un pendolo si calcola dalla formula

$$T = 2\pi\sqrt{\frac{l}{g}},$$

dove l è la sua lunghezza e g è l'accelerazione di gravità. Supponiamo che l e g siano noti con errori rispettivamente pari a Δl e Δg. Quindi l'errore nel calcolo di T sarà pari all'incremento ΔT corrispondente agli incrementi degli argomenti Δl e Δg. Sostituendo ΔT approssimativamente con dT, avremo

$$\Delta T \approx dT = \pi\left(\frac{\Delta l}{\sqrt{lg}} - \frac{\sqrt{l}\,\Delta g}{\sqrt{g^3}}\right).$$

I segni di Δl e Δg sono sconosciuti, ma possiamo ovviamente stimare ΔT in base alla disuguaglianza

$$|\Delta T| \leqslant \pi\left(\frac{|\Delta l|}{\sqrt{lg}} + \sqrt{\frac{l}{g^3}}\,|\Delta g|\right),$$

da cui dopo la divisione per T si ottiene

$$\frac{|\Delta T|}{T} < \left(\frac{|\Delta l|}{l} + \frac{|\Delta g|}{g}\right).$$

Possiamo quindi considerare in pratica che l'errore relativo per T è pari alla somma degli errori relativi per l e g.

Per simmetria di notazione, gli incrementi delle variabili indipendenti Δx e Δy sono solitamente indicati con i simboli dx e dy e sono anche chiamati differenziali. Con questa notazione il differenziale della funzione $u = f(x, y, z)$ può essere scritto così:

$$du = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz.$$

Le derivate parziali giocano un ruolo importante ogni volta che abbiamo a che fare con funzioni di più variabili, come accade in molte applicazioni dell'analisi alla tecnologia e alla fisica. Nel Capitolo VI affronteremo il problema di ricostruire una funzione a partire dalle proprietà delle sue derivate parziali.

Nei paragrafi seguenti diamo alcuni semplici esempi di applicazioni delle derivate parziali in analisi.

Differenziazione di funzioni implicite. Supponiamo di voler trovare la derivata di y, dove y è una funzione di x definita implicitamente dalla relazione

$$F(x, y) = 0 \qquad (38)$$

tra queste variabili. Se x e y soddisfano la relazione (38) e diamo a x l'incremento Δx, allora y riceverà un incremento Δy tale che $x + \Delta x$ e $y + \Delta y$ soddisfa nuovamente (38). Di conseguenza[30]

$$F(x + \Delta x, y + \Delta y) - F(x, y) = \frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y + \alpha\sqrt{\Delta x^2 + \Delta y^2} = 0.$$

Pertanto, posto che $\partial F/\partial y \neq 0$, ne consegue che

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = y'_x = -\frac{\dfrac{\partial F}{\partial x}}{\dfrac{\partial F}{\partial y}}.$$

In questo modo abbiamo ottenuto un metodo per trovare la derivata di una funzione implicita y senza prima risolvere l'equazione (38) per y.

Problemi massimi e minimi. Se una funzione, diciamo di due variabili $z = f(x, y)$, raggiunge il suo massimo nel punto (x_0, y_0), cioè se $f(x_0, y_0) \geqslant f(x, y)$ per tutti i punti (x, y) vicini a (x_0, y_0), allora questo punto deve essere anche il punto di massima altitudine per qualsiasi

linea formata dall'intersezione della superficie $z = f(x, y)$ con un piano parallelo a Oxz o Oyz. Quindi a questo punto dobbiamo avere

$$f'_x(x, y) = 0, f'_y(x, y) = 0. \qquad (39)$$

Le stesse equazioni devono valere anche per un punto di minimo locale. Di conseguenza, i valori maggiori o minimi della funzione vanno ricercati innanzitutto nei punti in cui sono soddisfatte le condizioni (39), ma non bisogna dimenticare nemmeno i punti sul confine del dominio di definizione della funzione e i punti in cui la funzione non ha derivata, se tali punti esistono.

Per stabilire se un punto (x, y) che soddisfa la (39) è effettivamente un punto di massimo o di minimo, si utilizzano spesso diversi argomenti indiretti. Ad esempio, se per qualsiasi motivo è chiaro che la funzione è differenziabile e raggiunge il suo minimo all'interno della regione e che c'è un solo punto in cui le condizioni (39) sono soddisfatte, allora ovviamente il minimo deve essere raggiunto in questo punto.

Ad esempio, si supponga di dover realizzare una scatola di latta rettangolare (senza coperchio) con volume assegnato V, utilizzando la minor quantità possibile di materiale. Se i lati della base di questa scatola sono indicati con x e y, allora la sua altezza h sarà uguale a V/xy, e di conseguenza la superficie S sarà data dalla funzione

$$S = xy + \frac{V}{xy}(2x + 2y) = xy + 2V\left(\frac{1}{x} + \frac{1}{y}\right) \qquad (40)$$

di x e y. Poiché x e y secondo i termini del problema devono essere positivi, la questione è stata ridotta a trovare il minimo della funzione $S(x, y)$ per tutti i possibili punti (x, y) nel primo quadrante del piano (x, y), che denoteremo con la lettera G.

Se il minimo viene raggiunto in qualche punto della regione G, allora le derivate parziali devono essere uguali a zero

$$\frac{\partial S}{\partial x} = y - \frac{2V}{x^2} = 0,$$

$$\frac{\partial S}{\partial y} = x - \frac{2V}{y^2} = 0,$$

ovvero $yx^2 = 2V$, $xy^2 = 2V$, da cui ricaviamo come dimensioni della scatola:

$$x = y = \sqrt[3]{2V} \quad \text{and} \quad h = \sqrt[3]{\frac{V}{4}}. \qquad (41)$$

Abbiamo risolto il problema ma non abbiamo dimostrato del tutto che la nostra soluzione sia corretta. Un matematico rigoroso ci dirà: "Avete supposto fin dall'inizio che nelle condizioni date la scatola di superficie minima esista effettivamente e, partendo da questa assunzione, ne avete trovate le dimensioni. Quindi in realtà avete ottenuto solo il seguente risultato: Se esiste un punto (x, y) in G per il quale la funzione S raggiunge il suo minimo, allora il le coordinate

di questo punto devono necessariamente essere determinate dall'equazione (41). Ma ora devi dimostrare che il minimo di S esiste per qualche punto in G e quindi I ammetterà la correttezza del tuo risultato. Questa osservazione è molto ragionevole, poiché, ad esempio, la nostra funzione S, come vedremo tra poco, non possiede alcun massimo nella regione G. Ma mostriamo come sia possibile convincersi che nel caso considerato la funzione effettivamente raggiunge il suo minimo in un certo punto (x, y) della regione G.

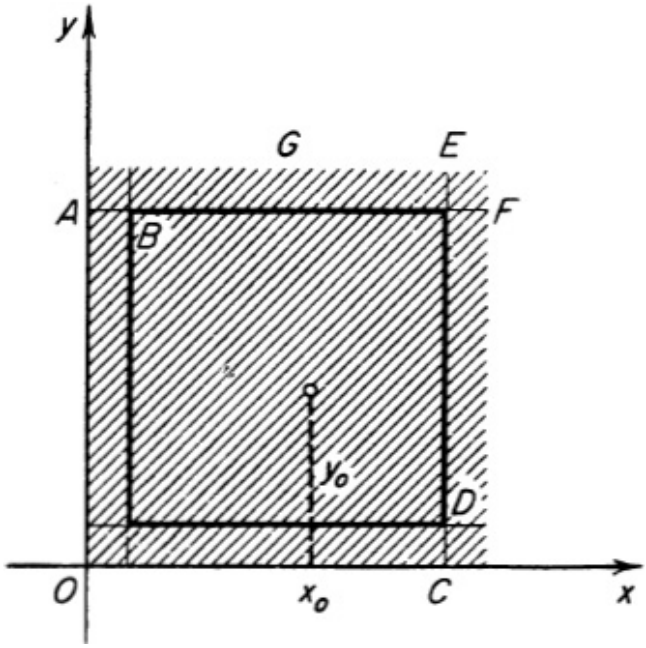


Figura 29.

Il teorema fondamentale su cui baseremo la nostra argomentazione è quello che viene dimostrato in analisi con assoluto rigore; equivale a quanto segue. Se la funzione f di una o più variabili è ovunque continua in una certa regione finita H che è limitata e comprende il suo confine, allora esiste sempre in H almeno un punto in cui la funzione raggiunge il suo minimo (massimo). Con questo teorema possiamo facilmente completare la nostra analisi del problema.

Consideriamo un punto arbitrario (x_0, y_0) della regione G; a questo punto poniamo $S(x_0, y_0) = N$. Scegliamo anche un numero R che soddisfi le due disuguaglianze $R > N$, $2VR > N$ e costruiamo un quadrato Ω_R, con lato R^2, come in figura 29, dove AB = CD = $1/R$.

Diamo ora un limite inferiore per i valori della nostra funzione $S(x, y)$ nei punti della regione G che si trovano all'esterno del quadrato Ω_R. Se il punto della regione G ha l'ascissa $x < 1/R$, allora

$$S(x, y) = xy + 2V\left(\frac{1}{x} + \frac{1}{y}\right) > 2V\frac{1}{x} > 2VR > N.$$

Analogamente, se il punto della regione G ha ordinata $y < 1/R$, allora anche $S > N$. Inoltre, se il punto della regione G ha l'ascissa $x > 1/R$ e se si trova al di sopra della retta AF oppure ha l'ordinata $y > 1/R$ e si trova a destra della retta CE, allora

$$S(x, y) > xy > \frac{1}{R} R^2 = R > N.$$

Pertanto, per tutti i punti (x, y) della regione G che si trovano all'esterno del quadrato Ω_R, vale la disuguaglianza $S(x, y) > N$, e poiché $S(x_0, y_0) = N$, il punto (x_0, y_0) deve appartenere al quadrato e di conseguenza il minimo della nostra funzione su G è pari al suo minimo sul quadrato.

Ma la funzione $S(x, y)$ è continua in questo quadrato e sul suo confine, sicché per il teorema sopra esposto esiste nel quadrato un punto (x, y) dove la nostra funzione assume il suo minimo per i punti del quadrato e di conseguenza per tutta la regione G. In tal modo è stata dimostrata l'esistenza di un minimo.

Questo argomento può servire da esempio del modo in cui è possibile discutere l'esistenza di un massimo o di un minimo per una funzione definita su un dominio illimitato.

La formula di Taylor. Come le funzioni di una variabile, le funzioni di più variabili possono essere rappresentate da una formula di Taylor. Ad esempio, un'espansione della funzione

$$u = f(x, y)$$

nell'intorno del punto (x_0, y_0) ha la forma seguente, se ci limitiamo alla prima e alla seconda potenza di $x - x_0$ e $y - y_0$:

$$\begin{aligned} f(x, y) = f(x_0, y_0) &+ [f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0)] \\ &+ \frac{1}{2!}[f''_{xx}(x_0, y_0)(x - x_0)^2 + 2f''_{xy}(x_0, y_0)(x - x_0)(y - y_0) \\ &+ f''_{yy}(x_0, y_0)(y - y_0)^2] + R_3 . \end{aligned}$$

Se la funzione $f(x, y)$ ha derivate parziali continue del secondo ordine, il termine resto qui si avvicinerà a zero più velocemente di

$$r^2 = (x - x_0)^2 + (y - y_0)^2,$$

cioè più veloce del quadrato della distanza tra i punti (x, y) e (x_0, y_0), come $r \to o$. La formula di Taylor fornisce un metodo ampiamente utilizzato per definire e calcolare approssimativamente i valori di varie funzioni.

Notiamo che con l'aiuto di questa formula possiamo anche rispondere alla domanda posta in precedenza, se una data funzione ha effettivamente un massimo o un minimo nel punto in cui $\partial f/\partial y = \partial f/\partial y = 0$. Infatti, se queste condizioni sono soddisfatte nel punto (x_0, y_0), quindi per i punti (x, y) vicini a (x_0, y_0), il valore della funzione sarà, secondo la formula di Taylor, diverso da $f(x_0, y_0)$ per l'importo

$$f(x, y) - f(x_0, y_0)$$

$$= \frac{1}{2!}[A(x - x_0)^2 + 2B(x - x_0)(y - y_0) + C(y - y_0)^2] + R_3, \qquad (42)$$

dove A, B e C indicano rispettivamente le derivate parziali seconde f''_{xx}, f''_{xy}, f''_{yy}, nel punto (x_0, y_0).

Se si scopre che la funzione

$$\Phi(x, y) = A(x - x_0)^2 + 2B(x - x_0)(y - y_0) + C(y - y_0)^2$$

è positivo per valori arbitrari di ($x - x_0$) e ($y - y_0$) non entrambi uguali a zero, allora anche il lato destro dell'equazione (42) sarà positivo per piccoli valori di ($x - x_0$) e ($y - y_0$), poiché per sufficientemente piccoli ($x - x_0$) e ($y - y_0$) la quantità R_3, è notoriamente inferiore in valore assoluto a $P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2), \cdots, P_n(x_n, y_n, z_n),$. Ne conseguirà quindi che nel punto (x_0, y_0) la funzione f raggiunge il suo minimo. D'altra parte, se la funzione $\Phi(x, y)$ è negativa per arbitrario ($x - x_0$) e ($y - y_0$) il lato destro di (42) sarà negativo per ($x - x_0$) e ($y - y_0$), così che nel punto (x_0, y_0) la funzione avrà un massimo. Nei casi più complicati è necessario considerare i termini successivi della formula di Taylor.

In modo del tutto analogo si possono trattare i problemi riguardanti il massimo o il minimo di funzioni di tre o più variabili. Come esercizio il lettore può dimostrarlo se vengono date le masse

$$m_1, m_2, \cdots, m_n$$

sono disposti nello spazio in determinati punti

$$P_1(x_1, y_1, z_1), P_2(x_2, y_2, z_2), \cdots, P_n(x_n, y_n, z_n),$$

il momento (d'inerzia) M di questo sistema di masse attorno al punto $P(x, y, z)$, definito come la somma dei prodotti delle masse e dei quadrati delle loro distanze dal punto P,

$$M(x, y, z) = \sum_{i=1}^{n} m_i[(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2],$$

sarà minimo se il punto P si trova nel cosiddetto baricentro del sistema, con le coordinate

$$x = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i}, \; y = \frac{\sum_{i=1}^{n} m_i y_i}{\sum_{i=1}^{n} m_i}, \; z = \frac{\sum_{i=1}^{n} m_i z_i}{\sum_{i=1}^{n} m_i}.$$

Massimi e minimi con condizioni sussidiarie. Per funzioni di più variabili possiamo porre diversi problemi riguardanti il massimo e il minimo. Illustriamolo con un semplice esempio. Supponiamo di voler trovare tra tutti i rettangoli inscritti in una circonferenza di raggio R quello di area maggiore. L'area di un rettangolo è uguale al prodotto xy dei suoi lati, dove x e y sono numeri positivi legati in questo caso dalla relazione $x^2 + y^2 = (2R)^2$, come risulta chiaro

dalla figura 30. Dobbiamo quindi trovare il massimo della funzione $f(x, y) = xy$ per tutti i x e y che soddisfano la relazione $x^2 + y^2 = 4R^2$.

Problemi di questo tipo, dove è necessario trovare il massimo (o il minimo) di una funzione $f(x, y)$ solo per quei valori di x e y che soddisfano una certa relazione che $\phi(x, y)$ sono molto comuni in pratica.

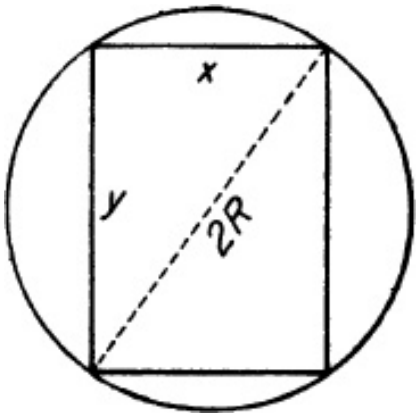


Figura 30.

Naturalmente sarebbe possibile risolvere l'equazione $\phi(x, y)$ per y, per sostituire la soluzione nella funzione $f(x, y)$ e cercare in questo modo il massimo ordinario per una funzione di una variabile x. Ma questo metodo è solitamente complicato e talvolta impossibile.

Per la soluzione di tali problemi in analisi è stata elaborata una procedura molto più conveniente chiamata metodo dei moltiplicatori di Lagrange. L'idea alla base è estremamente semplice. Consideriamo la funzione

$$F(x, y) = f(x, y) + \lambda\phi(x, y),$$

dove λ è un numero positivo arbitrario. Ovviamente, per x, y che soddisfa la condizione $\phi(x, y) = 0$, i valori di $F(x, y)$ coincidono con quelli di $f(x, y)$.

Per la funzione $F(x, y)$ cerchiamo un massimo senza condizioni di alcun genere su x e y. Nel punto massimo devono valere le condizioni $\partial F/\partial x = \partial F/\partial y = 0$[31]; in altre parole

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} = 0; \qquad (43)$$

$$\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} = 0. \qquad (44)$$

I valori di x e y nel punto massimo per $F(x, y)$, essendo una soluzione del sistema (43) e (44), dipendono dal coefficiente λ in queste equazioni. Supponiamo ora di essere riusciti a scegliere il numero λ in modo tale che le coordinate del punto massimo soddisfino la condizione

$$\phi(x, y) = 0. \qquad (45)$$

Allora questo punto sarà un massimo locale esatto per il problema originale.

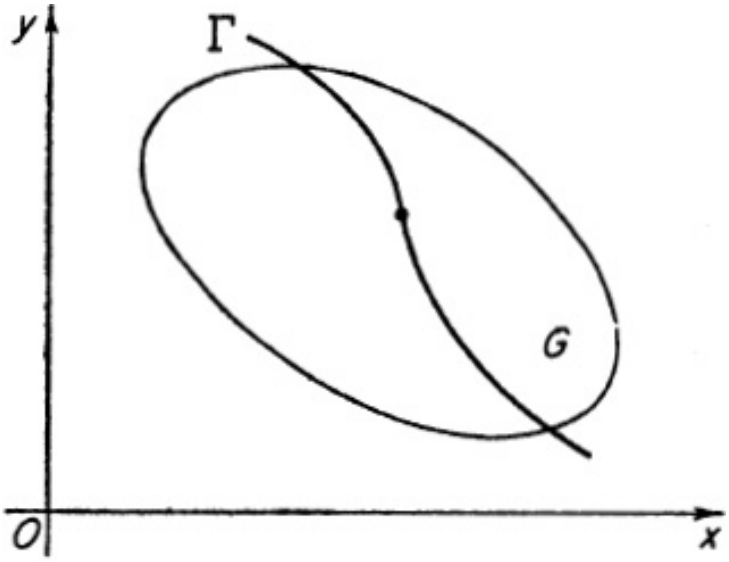


Figura 31.

In effetti, possiamo considerare il problema geometricamente nel modo seguente. La funzione $f(x, y)$ è definita su una determinata regione G (figura 31). La condizione $\phi(x, y) = 0$ sarà normalmente soddisfatta dai punti di una curva Γ. Dobbiamo trovare il valore massimo di $f(x, y)$ sui punti della curva Γ. Se $F(x, y)$ raggiunge il suo massimo sulla curva Γ, allora $F(x, y)$ non aumenta per piccoli spostamenti in una direzione arbitraria da questo punto, e in particolare per spostamenti lungo la curva Γ. Ma per spostamenti lungo Γ, i valori di $F(x, y)$, coincidono con quelli di $f(x, y)$ il che significa che per piccoli spostamenti lungo la curva la funzione $f(x, y)$ non aumenta, o in altre parole ha il massimo locale nel punto.[E1]

Questi argomenti indicano un metodo semplice per risolvere il problema. Risolviamo le equazioni (43), (44), (45) per le incognite x, y e λ, ottenendo una o più soluzioni

$$(x_1, y_1, \lambda_1),\ (x_2, y_2, \lambda_2),\ \cdots . \qquad (46)$$

Ai punti (x_1, y_1), (x_2, y_2), ⋯ così determinati uniamo quei punti del confine di G dove la curva Γ lascia la regione G. Quindi tra tutti questi punti scegliamo quello in cui $f(x, y)$ assume il suo valore massimo (o minimo).

Naturalmente, le argomentazioni qui esposte sono lungi dal dimostrare la correttezza del metodo. Infatti, non abbiamo ancora nemmeno dimostrato che i punti di massimo locale per $f(x, y)$ sulla curva Γ possano essere ottenuti come punti di massimo per la funzione $F(x, y)$ per qualche valore di λ. Tuttavia è possibile dimostrare, come fatto nei libri di testo in analisi, che ogni punto (x_0, y_0) dove $f(x, y)$ ha un massimo locale sarà ottenuto con il metodo indicato, a patto solo che in questo punto puntare le derivate parziali $\phi'_x(x_0, y_0)$ e $\phi'_y(x_0, y_0)$ non sono entrambi uguali a zero.[32]

Usiamo il metodo di Lagrange per risolvere il problema all'inizio della presente sezione. In questo caso $f(x, y) = xy$; $\phi(x, y) = x^2 + y^2 - 4R^2$. Impostiamo le equazioni (43), (44), (45)

$$y + 2\lambda x = 0,$$
$$x + 2\lambda y = 0,$$
$$x^2 + y^2 = 4R^2,$$

per cui, tenendo conto che x e y sono positivi, troviamo l'unica soluzione

$$x = y = R\sqrt{2}\left(\lambda = -\frac{1}{2}\right).$$

Per questi valori di x e y, che sono uguali tra loro affinché il rettangolo inscritto sia un quadrato, l'area è infatti massima.

Il metodo di Lagrange può essere esteso per trattare funzioni di tre o più variabili. Potrebbe esserci un numero qualsiasi di condizioni sussidiarie (minori del numero di variabili) del tipo di condizione (45) e introdurremo il numero corrispondente di moltiplicatori ausiliari.

Diamo alcuni esempi di problemi che coinvolgono massimi o minimi con condizioni sussidiarie.

Esempio 1. Per quale altezza h e raggio r un serbatoio cilindrico aperto di un dato volume V richiederà la minima quantità di lamiera per la sua fabbricazione; cioè l'area dei suoi lati e della base circolare sarà minima ?

Il problema ovviamente si riduce a trovare il minimo della funzione delle variabili r e h

$$f(r, h) = 2\pi rh + \pi r^2$$

nella condizione $\pi r^2 h = V$, che può essere scritta nella forma

$$\phi(r, h) = \pi r^2 h - V = 0.$$

Esempio 2. È necessario che un punto in movimento passi da A a B (figura 32). Sul percorso AM si muove con la velocità v_1, e su MB con la velocità v_2. Dove va posizionato il punto M sulla linea DD′ affinché l'intero percorso da A a B possa essere percorso il più rapidamente possibile?

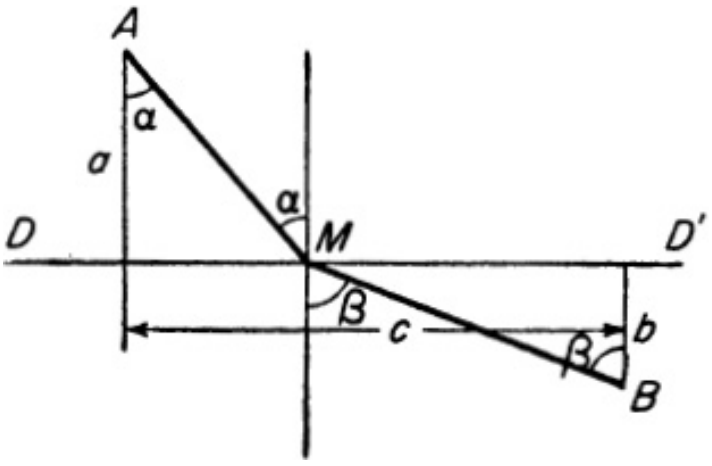


Figura 32.

Prendiamo come incogniti gli angoli α e β segnati in figura 32. Sono note le lunghezze a e b delle perpendicolari dai punti A e B alla retta DD′ e la distanza c tra loro. Il tempo necessario per percorrere l'intero percorso è rappresentato, come si può facilmente vedere, dalla formula

$$f(\alpha, \beta) = \frac{a}{v_1 \cos \alpha} + \frac{b}{v_2 \cos \beta}.$$

È necessario trovare il minimo di questa espressione, tenendo conto del fatto che α e β sono collegati dalla relazione

$$a \tan \alpha + b \tan \beta = c.$$

Il lettore può risolvere questi esempi con il metodo di Lagrange. Nel secondo esempio troverà che la posizione migliore per M è data dalla condizione

$$\frac{\sin \alpha}{\sin \beta} = \frac{v_1}{v_2}.$$

Questa è la ben nota legge della rifrazione della luce. Di conseguenza un raggio di luce sarà rifratto nel suo passaggio da un mezzo all'altro in modo tale che il tempo del suo passaggio da un punto dell'uno al punto dell'altro sia minimo. Conclusioni di questo tipo sono interessanti non solo ai fini computazionali ma anche da un punto di vista filosofico generale; hanno ispirato i ricercatori delle scienze esatte a penetrare sempre più nelle leggi profonde e generali della natura.

Notiamo infine che i moltiplicatori λ, introdotti nella soluzione dei problemi con il metodo di Lagrange, non sono meri numeri ausiliari. In ogni caso sono strettamente collegati alla natura essenziale del particolare problema e hanno un'interpretazione concreta.

§ 13. Generalizzazioni del concetto di integrale

Nel §10 abbiamo definito l'integrale definito della funzione $f(x)$ sull'intervallo $[a, b]$ come limite della somma

$$\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i$$

quando la lunghezza del segmento maggiore Δx_i nella suddivisione di $[a, b]$ si avvicina a zero. Nonostante la classe di funzioni $f(x)$ per la quale esiste effettivamente questo limite (la classe delle funzioni integrabili) sia molto ampia, e in particolare comprenda tutte le funzioni continue e anche molte discontinue, questa classe di funzioni presenta un grave difetto. Se aggiungiamo, sottraiamo, moltiplichiamo o in determinate condizioni dividiamo i valori di due funzioni integrabili $f(x)$ e $\phi(x)$, otteniamo funzioni che, come si può facilmente dimostrare, sono nuovamente integrabili. Per $f(x)/\phi(x)$ questo sarà vero in tutti i casi in cui $1/\phi(x)$ rimane limitato su $[a, b]$. Ma se una funzione si ottiene come risultato di un passaggio al limite da una successione di funzioni integrabili approssimate $f_1(x), f_2(x), f_3(x), \cdots$ tale che per tutti i valori di x nell'intervallo $[a, b]$

$$f(x) = \lim_{n\to\infty} f_n(x),$$

allora la funzione limite $f(x)$ non è necessariamente integrabile.

In molti casi questa ed altre circostanze danno origine a notevoli complicazioni, poiché il processo di passaggio al limite è ampiamente utilizzato.

Una via d'uscita dalla difficoltà è stata scoperta effettuando ulteriori generalizzazioni del concetto di integrale. Il più importante di questi è l'integrale di Lebesgue, che il lettore conoscerà nel capitolo XV sulla teoria delle funzioni di variabile reale. Ma qui ci limiteremo alle generalizzazioni dell'integrale in altre direzioni, che sono anche della massima importanza nella pratica.

Integrali multipli. Abbiamo già studiato il processo di integrazione per funzioni di una variabile definita su una regione unidimensionale, vale a dire un intervallo. Ma il processo analogo può essere esteso a funzioni di due, tre o più variabili, definite su regioni corrispondenti.

Consideriamo ad esempio una superficie

$$z = f(x, y)$$

definita in un sistema di coordinate rettangolare, e sul piano Oxy sia data una regione G delimitata da una curva chiusa Γ. Occorre trovare il volume delimitato dalla superficie, dal piano Oxy e dalla superficie cilindrica passante per la curva Γ con generatrici parallele all'asse Oz (figura 33). Per risolvere questo problema dividiamo la regione piana G in sottoregioni mediante una rete di rette parallele agli assi Ox e Oy e denotiamo con

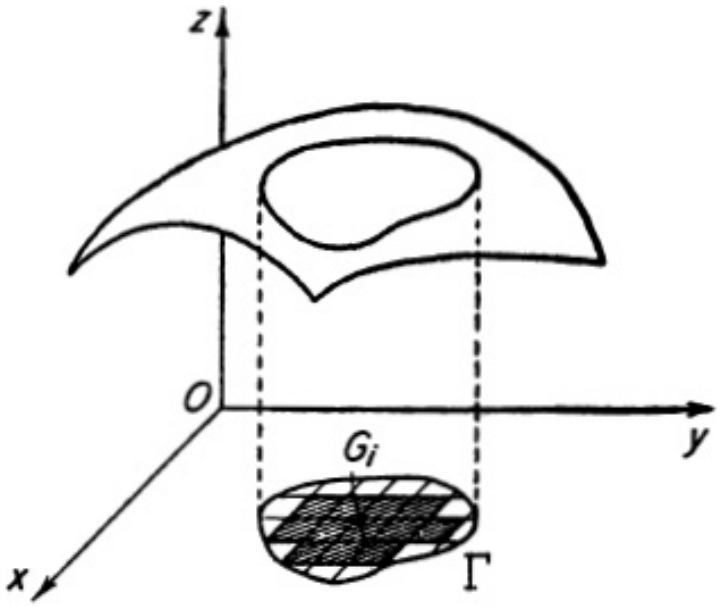


Figura 33.

$$G_1 , G_2 , \cdots, G_n$$

quelle sottoregioni costituite da rettangoli completi. Se la rete è sufficientemente fine, praticamente tutta la regione *G* sarà coperta dai rettangoli enumerati. In ognuno di essi scegliamo a piacimento un punto

$$(\xi_1 , \eta_1), (\xi_2 , \eta_2), \cdots, (\xi_n , \eta_n)$$

e, assumendo per semplicità che G_i indichi non solo il rettangolo ma anche la sua area, impostiamo la somma

$$S_n = f(\xi_1 , \eta_1)\, G_1 + f(\xi_2 , \eta_2)\, G_2 + \cdots + f(\xi_n , \eta_n)\, G_n = \sum_{i=1}^{n} f(\xi_i , \eta_i)\, G_i . \tag{47}$$

È chiaro che, se la superficie è continua e la rete è sufficientemente fine, questa somma può avvicinarsi quanto si vuole al volume desiderato *V*. Otterremo il volume desiderato esattamente se prendiamo il limite della somma (47) per suddivisioni sempre più fini (cioè per suddivisioni tali che la maggiore delle diagonali dei nostri rettangoli si avvicini allo zero)

$$\lim_{\max d(G_i)\to 0} \sum_{i=1}^{n} f(\xi_i , \eta_i)\, G_i = V. \tag{48}$$

Dal punto di vista dell'analisi è quindi necessario, per poter determinare il volume *V*, effettuare una certa operazione matematica sulla funzione $f(x, y)$ e sul suo dominio di definizione *G*, operazione indicata dal membro sinistro dell'uguaglianza (48). Questa operazione è chiamata integrazione della funzione *f* sulla regione *G* e il suo risultato è l'integrale di *f* su *G*. È consuetudine denotare questo risultato nel modo seguente:

$$\iint_G f(x, y)\, dx\, dy = \lim_{\max d(G_i)\to 0} \sum_{i=1}^{n} f(\xi_i , \eta_i)\, G_i . \qquad (49)$$

Allo stesso modo, possiamo definire l'integrale di una funzione di tre variabili su una regione tridimensionale *G*, che rappresenta un certo corpo nello spazio. Ancora una volta dividiamo la regione *G* in parti, questa volta per piani paralleli ai piani delle coordinate. Tra queste parti scegliamo quelle che rappresentano parallelepipedi completi e le enumeriamo

$$G_1 , G_2 , \cdots, G_n .$$

In ognuno di questi scegliamo un punto arbitrario

$$(\xi_1 , \eta_1 , \zeta_1), (\xi_2 , \eta_2 , \zeta_2), \cdots, (\xi_n , \eta_n , \zeta_n)$$

e formiamo la somma

$$S = \sum_{i=1}^{n} f(\xi_i , \eta_i , \zeta_i)\, G_i , \qquad (50)$$

dove G_i indica il volume del parallelepipedo G_i. Infine definiamo l'integrale di $f(x, y, z)$ sulla regione *G* come limite

$$\lim_{\max d(G_i)\to 0} \sum_{i=1}^{n} f(\xi_i , \eta_i , \zeta_i)\, G_i = \iiint_G f(x, y, z)\, dx\, dy\, dz, \qquad (51)$$

a cui tende la somma (50) quando la diagonale maggiore $d(G_i)$ si avvicina a zero.

Consideriamo un esempio. Immaginiamo che la regione *G* sia riempita da una massa non omogenea la cui densità in ogni punto di *G* è data da una funzione nota $\rho(x, y, z)$. La densità $\rho(x, y, z)$ della massa nel punto (x, y, z) è definita come il limite avvicinato dal rapporto tra la massa di una piccola regione arbitraria contenente il punto (x, y, z) al volume di questa regione man mano che il suo diametro si avvicina allo zero.[33] Per determinare la massa del corpo *G* è naturale procedere come segue. Dividiamo la regione *G* in parti per piani paralleli ai piani delle coordinate ed enumeriamo i parallelepipedi completi così formati

$$G_1 , G_2 , \cdots, G_n .$$

Supponendo che i piani divisori siano sufficientemente vicini tra loro, commetteremo solo un piccolo errore se trascuriamo le regioni irregolari del corpo e definiamo la massa di ciascuna delle regioni regolari G_i (i parallelepipedi completi) come il prodotto

$$\rho(\xi_i , \eta_i , \zeta_i) G_i ,$$

dove (ξ_i, η_i, ζ_i) è un punto arbitrario G_i. Di conseguenza il valore approssimativo della massa *M* sarà espresso dalla somma

$$S_n = \sum_{i=1}^{n} \rho(\xi_i , \eta_i , \zeta_i)\, G_i ,$$

ed il suo valore esatto sarà chiaramente il limite di tale somma man mano che la diagonale maggiore G_i si avvicina allo zero; questo è

$$M = \iiint_G \rho(x, y, z)\, dx\, dy\, dz = \lim_{\max d(G_i)\to 0} \sum_{i=1}^{n} \rho(\xi_i , \eta_i , \zeta_i)\, G_i .$$

Gli integrali (49) e (51) sono chiamati rispettivamente integrali doppi e tripli.

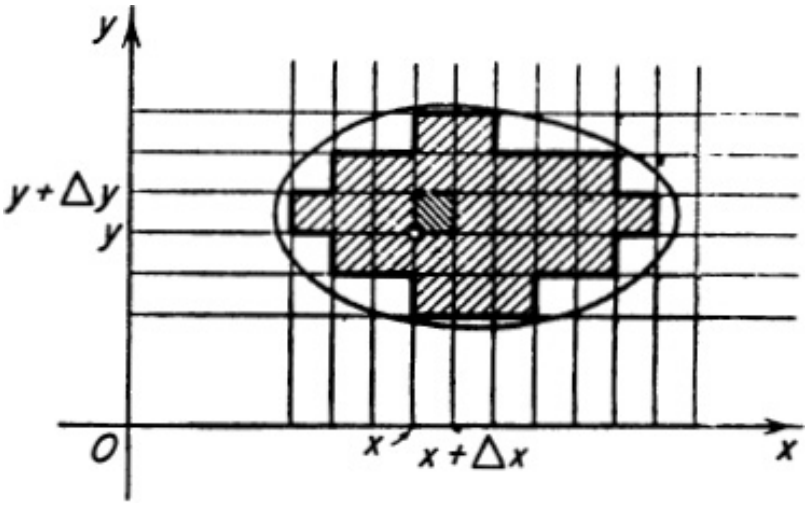


Figura 34.

Esaminiamo un problema che porta ad un integrale doppio. Immaginiamo che l'acqua scorra su una superficie piana. Inoltre, su questa superficie l'acqua sotterranea filtra (o ritorna nel terreno) con un'intensità $f(x, y)$ che è diversa in punti diversi. Consideriamo una regione G delimitata da un contorno chiuso (figura 34) e assumiamo che in ogni punto di G conosciamo l'intensità $f(x, y)$, ovvero la quantità di acqua sotterranea che filtra al minuto per cm^2 di superficie; avremo $f(x, y) > 0$ dove l'acqua filtra e $f(x, y) < 0$ dove penetra nel terreno. Quanta acqua si accumulerà sulla superficie G al minuto?

Se dividiamo G in piccole parti, consideriamo la velocità di infiltrazione approssimativamente costante in ciascuna parte e poi passiamo al limite per suddivisioni sempre più fini, otterremo un'espressione per tutta la quantità di acqua accumulata sotto forma di un integrale

$$\iint_G f(x, y)\, dx\, dy.$$

Gli integrali doppi (duplici) furono introdotti per la prima volta da Eulero. Gli integrali multipli costituiscono uno strumento che viene utilizzato tutti i giorni in calcoli e indagini dei più svariati generi.

Sarebbe anche possibile mostrare, anche se non lo faremo qui, che il calcolo di integrali multipli può essere ridotto, di regola, al calcolo iterato di integrali unidimensionali ordinari.

Integrali di contorno e di superficie. Dobbiamo infine menzionare che sono possibili ancora altre generalizzazioni dell'integrale. Ad esempio, il problema di definire il lavoro compiuto da una forza variabile applicata a un punto materiale, mentre quest'ultimo si muove lungo una data curva, porta naturalmente a un cosiddetto integrale curvilineo, e il problema di trovare la carica totale su una superficie su cui l'elettricità è distribuita continuamente con una data densità superficiale porta a un altro nuovo concetto, un integrale su una superficie curva.

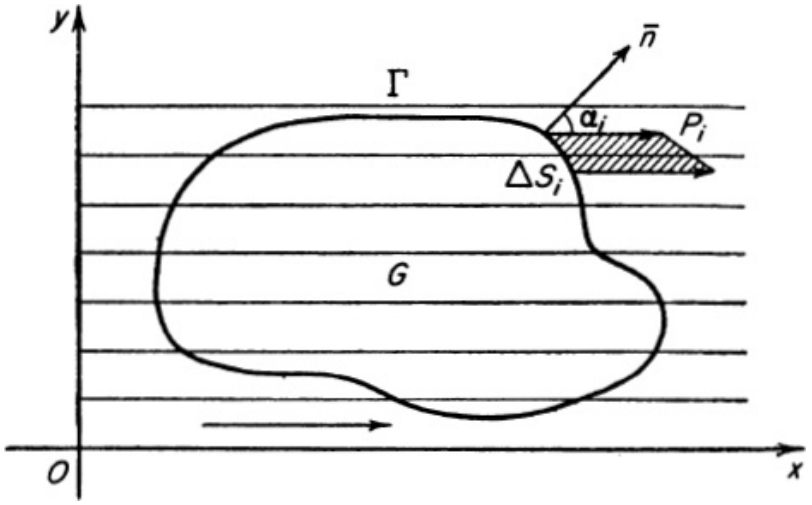


Figura 35.

Ad esempio, supponiamo che un liquido fluisca nello spazio (figura 35) e che la velocità di una particella del liquido nel punto (x, y) sia data da una funzione $P(x, y)$, non dipendente da z. Se desideriamo determinare la quantità di liquido che scorre al minuto attraverso il contorno Γ,[34] possiamo ragionare nel modo seguente. Dividiamo Γ in segmenti Δs_i. La quantità di acqua che scorre attraverso un segmento Δs_i è approssimativamente uguale alla colonna di liquido evidenziata in figura 35; questa colonna può essere considerata come la quantità di liquido che si fa strada al minuto attraverso quel segmento del contorno. Ma l'area del parallelogramma ombreggiato è uguale a

$$P_i(x, y) \cdot \Delta s_i \cdot \cos \alpha_i ,$$

dove α_i è l'angolo tra la direzione $\bar{x}$ dell'asse x e la normale esterna della superficie delimitata dal contorno Γ; questa normale è la perpendicolare $\bar{n}$ alla tangente, che possiamo considerare come definente la direzione del segmento Δs_i. Sommando le aree di tali parallelogrammi e passando al limite per suddivisioni sempre più fini del contorno Γ, determiniamo la quantità di acqua che scorre al minuto attraverso il contorno Γ; viene indicato così:

$$\int_{\Gamma} P(x, y) \cos(\bar{n}, \bar{x})\, ds$$

ed è detto integrale curvilineo. Se il flusso non è parallelo ovunque, la sua velocità in ciascun punto (x, y) avrà una componente $P(x, y)$ lungo l'asse x e una componente $Q(x, y)$ lungo l'asse y. In questo caso possiamo dimostrare con un ragionamento analogo che la quantità di acqua che scorre attraverso il contorno sarà pari a

$$\int_{\Gamma} [P(x, y) \cos(\bar{n}, \bar{x}) + Q(x, y) \cos(\bar{n}, \bar{y})]\, ds.*$$

Per la forma equivalente dell'integrale, si veda la nota [35].

Quando parliamo di integrale su una superficie curva G per una funzione $f(M)$ dei suoi punti $M(x, y, z)$, intendiamo il limite delle somme della forma

$$\lim \sum_{i=1}^{n} f(M_i)\, \Delta\sigma_i = \iint_{G} f(x, y, z)\, d\sigma$$

per suddivisioni sempre più fini della regione G in porzioni le cui aree sono uguali a $\Delta\sigma_i$.

Esistono metodi generali per trasformare integrali multipli, curvilinei e di superficie in altre forme e per calcolare i loro valori, esattamente o approssimativamente.

Formula di Ostrogradskij. Diverse formule importanti e molto generali che mettono in relazione un integrale su un volume con un integrale sulla sua superficie (e anche un integrale su una superficie, curva o piana, con un integrale attorno al suo confine) furono scoperte a metà del secolo scorso da Ostrogradskij.

Non cercheremo di dare qui una dimostrazione della formula generale di Ostrogradskij, che ha un'applicazione molto ampia, ma la illustreremo semplicemente con un esempio del suo caso particolare più semplice.

Immaginiamo, come abbiamo fatto prima, che su una superficie piana si svolga un flusso orizzontale d'acqua che anch'essa penetra nel terreno o ne fuoriesce nuovamente. Delineiamo una regione G, delimitata da una curva Γ, e assumiamo che per ogni punto della regione conosciamo le componenti $P(x, y)$ e $Q(x, y)$ della velocità dell'acqua rispettivamente nella direzione dell'asse x e dell'asse y.

Calcoliamo la velocità con cui l'acqua filtra dal suolo in un punto con coordinate (x, y). A questo scopo consideriamo un piccolo rettangolo di lati Δx e Δy situato nel punto (x, y).

Come risultato della velocità $P(x, y)$ attraverso il bordo verticale sinistro di questo rettangolo, scorreranno circa $P(x, y)\Delta y$ unità d'acqua al minuto nel rettangolo e attraverso il lato destro nello stesso tempo usciranno circa Unità $P(x + \Delta x, y)\Delta y$. In generale, la quantità netta di acqua che lascia un'unità quadrata di superficie come risultato del flusso attraverso i suoi lati verticali sinistro e destro sarà approssimativamente

$$\frac{[P(x + \Delta x, y) - P(x, y)]\, \Delta y}{\Delta x\, \Delta y}.$$

Se lasciamo che Δx si avvicini allo zero, otteniamo il limite

$$\frac{\partial P}{\partial x}.$$

Corrispondentemente, la portata netta dell'acqua per unità di area nella direzione dell'asse y sarà data da

$$\frac{\partial Q}{\partial y}.$$

Ciò significa che l'intensità della infiltrazione delle acque sotterranee nel punto con le coordinate (x, y) sarà pari a

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}.$$

Ma in generale, come abbiamo visto prima, la quantità di acqua in uscita dal suolo sarà data dall'integrale doppio della funzione che esprime l'intensità della infiltrazione di acqua sotterranea in ciascun punto, ovvero

$$\iint_G \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dx\, dy. \qquad (52)$$

Ma poiché l'acqua è incomprimibile, tutta questa quantità deve defluire nello stesso tempo attraverso i limiti del contorno Γ. La quantità di acqua che defluisce dal contorno Γ è espressa, come abbiamo visto prima, dall'integrale curvilineo su Γ

$$\int_\Gamma [P(x, y)\cos(\bar{n}, \bar{x}) + Q(x, y)\cos(\bar{n}, \bar{y})]\, ds. \qquad (53)$$

L'uguaglianza delle grandezze (52) e (53) esprime la formula di Ostrogradskij nel suo caso bidimensionale più semplice

$$\iint_G \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dx\, dy = \int_\Gamma [P(x, y)\cos(\bar{n}, \bar{x}) + Q(x, y)\cos(\bar{n}, \bar{y})]\, ds.$$

Abbiamo semplicemente spiegato il significato di questa formula con un esempio fisico, ma può essere dimostrato matematicamente.

In questo modo il teorema matematico di Ostrogradskij riflette un fenomeno diffuso nel mondo esterno, che nel nostro esempio abbiamo interpretato in modo facilmente visualizzabile come conservazione del volume di un fluido incomprimibile.

Ostrogradskij stabilì una formula molto più generale che esprime la connessione tra un integrale su un volume multidimensionale e un integrale sulla sua superficie. In particolare, per un corpo tridimensionale G, delimitato dalla superficie Γ, la sua formula è

$$\iiint_G \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dx\, dy\, dz$$
$$= \iint_\Gamma [P \cos(\bar{n}, \bar{x}) + Q \cos(\bar{n}, \bar{y}) + R \cos(\bar{n}, \bar{z})]\, d\sigma,$$

dove $d\sigma$ è l'elemento di superficie.

È interessante notare che la formula fondamentale del calcolo integrale

$$\int_a^b f(x)\, dx = F(b) - F(a) \qquad (54)$$

può essere considerato come un caso unidimensionale della formula di Ostrogradskij. L'equazione (54) collega l'integrale su un intervallo con l'"integrale" sul suo confine "nullo-dimensionale", costituito dai due punti finali.

La formula (54) può essere illustrata dalla seguente analogia. Immaginiamo che in un tubo diritto con sezione costante $s = 1$ l'acqua scorra con velocità $F(x)$, che è diversa per le diverse sezioni trasversali (figura 36). Attraverso le pareti porose del tubo, l'acqua penetra al suo interno (o fuori da esso) ad una velocità che è diversa anche per le diverse sezioni trasversali.

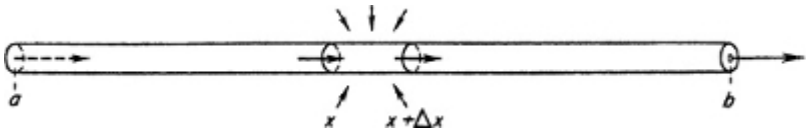


Figura 36.

Se consideriamo un tratto di tubo da x a $x + \Delta x$, la quantità di acqua che vi penetra nell'unità di tempo deve essere compensata dalla differenza $F(x + \Delta x) - F(x)$ tra la quantità che esce da questo segmento e la quantità che vi scorre lungo il tubo. Quindi la quantità che filtra nel segmento è pari alla differenza $F(x + \Delta x) - F(x)$, e di conseguenza il tasso di infiltrazione per unità di lunghezza del tubo (il rapporto tra la infiltrazione su un segmento infinitesimo e la lunghezza del segmento) sarà pari a

$$f(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x).$$

Più in generale, la quantità d'acqua che penetra nel tubo lungo tutta la sezione $[a, b]$ deve essere pari alla quantità persa per flusso alle estremità del tubo. Ma la quantità che filtra attraverso le pareti è pari a $\int_a^b f(x)\, dx$ e la quantità persa a causa del flusso attraverso le estremità è $F(b) - F(a)$. L'uguaglianza di queste due grandezze produce la formula (54).

§ 14. Serie

Concetto di serie. Una serie in matematica è un'espressione della forma

$$u_0 + u_1 + u_2 + \cdots.$$

I numeri u_k sono chiamati termini della serie. Essi sono infiniti e sono disposti in un ordine definito, per cui ad ogni numero naturale $k = 0, 1, 2, \cdots$ corrisponde un valore definito u_k.

Il lettore tenga presente che non abbiamo detto se sia possibile calcolare un valore per tali espressioni o, eventualmente, come farlo. La presenza di un segno più tra i termini u_k nella nostra espressione sembra indicare che in qualche modo tutti i termini dovrebbero essere sommati. Ma ce ne sono infiniti e l'addizione di numeri è definita solo per un numero finito di termini.

Indichiamo con S_n la somma dei primi n termini della serie; la chiameremo la n-esima somma parziale. Di conseguenza otteniamo una successione di numeri

$$\begin{aligned} S_1 &= u_0\,, \\ S_2 &= u_0 + u_1\,, \\ &\cdots\cdots\cdots\cdots\cdots \\ S_n &= u_0 + u_1 + \cdots + u_{n-1}\,, \\ &\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

e possiamo parlare di una quantità variabile S_n, dove $n = 1,2, \cdots$.

La serie si dice convergente se, quando $n \to \infty$, la variabile S_n si avvicina ad un limite finito definito

$$\lim_{n\to\infty} S_n = S.$$

Questo limite si chiama somma della serie, e in questo caso si scrive

$$S = u_0 + u_1 + u_2 + \cdots.$$

Ma se, quando $n \to \infty$, il limite S_n non esiste, allora la serie si dice divergente e in questo caso non ha senso parlare della sua somma. [36] Ma se tutti gli u_n hanno lo stesso segno, allora si è soliti dire che la somma della serie è uguale a infinito con il segno corrispondente.

Consideriamo ad esempio la serie

$$1 + x + x^2 + \cdots,$$

i cui termini formano una progressione geometrica con rapporto comune x.

La somma dei primi n termini è uguale a

$$S_n(x) = \frac{1 - x^n}{1 - x} (x \neq 1); \qquad (55)$$

se | x | < 1 questa somma ha un limite

$$\lim_{n \to \infty} S_n(x) = \frac{1}{1 - x},$$

e così per | x | < 1 possiamo scrivere

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

Se | x | > 1, allora ovviamente

$$\lim_{n \to \infty} S_n(x) = \infty,$$

e la serie diverge. La stessa situazione vale anche per x = 1, come si vede immediatamente senza l'uso della formula (55), che per x = 1 non ha significato. Infine, se x = -1 le somme parziali assumono alternativamente i valori +1 e 0, per cui anche questa serie è divergente.

Ad ogni serie corrisponde una determinata successione di valori delle sue somme parziali S_1, S_2, S_3, ⋯ tale che la convergenza della serie dipende dal fatto che le somme si avvicinano ad un limite. Viceversa, ad una successione arbitraria di numeri S_1, S_2, S_3, ⋯ corrisponde una serie

$$S_1 + (S_2 - S_1) + (S_3 - S_2) + \cdots,$$

le cui somme parziali saranno i numeri della successione. Quindi la teoria delle variabili che si estendono su una successione può essere ridotta alla teoria delle serie corrispondenti e viceversa. Eppure ciascuna di queste teorie ha un significato indipendente. In alcuni casi è più conveniente studiare direttamente la variabile ed in altri considerare la serie equivalente.

Notiamo che le serie sono servite per molto tempo come un metodo importante per rappresentare varie entità (soprattutto funzioni) e per calcolarne il valore. Naturalmente, le opinioni dei matematici riguardo alle serie sono cambiate con il passare del tempo, corrispondentemente ai cambiamenti nelle loro idee sugli infinitesimi. La chiara definizione di convergenza e divergenza di una serie è stata formulata all'inizio del secolo scorso contemporaneamente al concetto strettamente correlato di limite.

Se la serie converge, allora il suo termine generale si avvicina a zero all'aumentare di n

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

Dagli esempi forniti nei paragrafi seguenti, sarà chiaro che l'affermazione contraria è in generale falsa. Ma il criterio è comunque utile, poiché fornisce una condizione necessaria per la convergenza di una serie. Ad esempio, la divergenza di una progressione geometrica con

rapporto comune $x > 1$ consegue immediatamente dal fatto che il suo termine generale non si avvicina allo zero.

Se la serie è composta da termini positivi, la sua somma parziale S_n aumenta all'aumentare di n e possono esistere solo due casi: o la variabile S_n diventa e rimane maggiore di qualsiasi numero preassegnato A per n sufficientemente grande, nel qual caso $\lim_{n\to\infty} S_n = \infty$, per cui la serie diverge; oppure esiste un numero A tale che per tutti gli n il valore di S_n non supera A; ma allora la variabile S_n si avvicina necessariamente ad un limite finito definito non maggiore di A e la serie è convergente.

Convergenza di una serie. La questione se una data serie converge o diverge può spesso essere risolta confrontandola con un'altra serie. Qui è consuetudine avvalersi del seguente criterio.

Se ci vengono fornite due serie

$$u_0 + u_1 + u_2 + \cdots,$$
$$v_0 + v_1 + v_2 + \cdots$$

con termini positivi tali che per tutti i valori di n, a partire da un certo, abbiamo la disuguaglianza

$$u_n \leqslant v_n ,$$

allora la convergenza della seconda serie implica la convergenza della prima, e la divergenza della prima implica la divergenza della seconda.

Consideriamo ad esempio la cosiddetta serie armonica

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

I suoi termini non sono corrispondentemente inferiori ai termini della serie

$$1 + \underline{\frac{1}{2}} + \underline{\frac{1}{4} + \frac{1}{4}} + \underline{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}} + \underbrace{\frac{1}{16} + \cdots + \frac{1}{16}}_{8 \text{ times}} + \cdots ,$$

in cui la somma dei termini sottolineati è in ciascun caso pari a $\frac{1}{2}$.

È chiaro che la somma S_n della seconda serie tende all'infinito all'aumentare di n, e di conseguenza che la serie armonica diverge.

La serie

$$1 + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \frac{1}{4^\alpha} + \cdots , \qquad (56)$$

dove α è un numero positivo inferiore all'unità, ovviamente diverge, poiché per n arbitrario

$$\frac{1}{n^\alpha} > \frac{1}{n} (0 < \alpha < 1).$$

D'altra parte è possibile dimostrare che la serie (56) per $\alpha > 1$ è convergente. Lo dimostreremo qui solo per il caso $\alpha \leqslant 2$; a questo scopo consideriamo la serie

$$\left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) + \cdots$$

con termini positivi. Converge all'unità in quanto la sua somma, poiché le sue somme parziali S_n, sono pari a

$$S_n = 1 - \frac{1}{n+1} \to 1 \ (n \to \infty).$$

D'altra parte, il termine generale di questa serie soddisfa la disuguaglianza

$$\frac{1}{n-1} - \frac{1}{n} = \frac{1}{(n-1)\,n} > \frac{1}{n^2},$$

da cui segue che la serie

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots$$

converge. Tanto più allora la serie (56) convergerà con $\alpha > 2$.

Diamo qui senza dimostrazione un altro criterio utile per la convergenza e la divergenza delle serie a termini positivi, il cosiddetto criterio di d'Alembert.

Supponiamo che, man mano che n tende all'infinito, il rapporto $(u_n + 1)/u_n$ abbia limite q. Allora per $q < 1$ la successione convergerà sicuramente, mentre per $q > 1$ divergerà. Ma per $q = 1$ la questione della sua convergenza resta aperta.

La somma di un numero finito di addendi non cambia se permutiamo gli addendi. Ma in generale questo non è più vero per le serie infinite. Esistono serie convergenti per le quali è possibile permutare i termini in modo da cambiarne la somma e addirittura trasformarli in serie divergenti. Le serie con somme instabili di questo tipo non riescono a possedere una delle proprietà fondamentali delle somme ordinarie, la permutabilità degli addendi. Quindi è importante distinguere quelle serie che preservano questa proprietà. Si scopre che sono le cosiddette serie assolutamente convergenti. La serie

$$u_0 + u_1 + u_2 + u_3 + \cdots$$

si dice assolutamente convergente se la serie

$$|u_0| + |u_1| + |u_2| + |u_3| + \cdots$$

converge anche il numero dei valori assoluti dei suoi termini. È possibile dimostrare che una serie assolutamente convergente è sempre convergente; in altre parole, che le sue somme parziali S_n si avvicinano ad un limite finito. È ovvio che ogni serie convergente con termini di un segno è assolutamente convergente.

La serie

$$|u_0| + |u_1| + |u_2| + |u_3| + \cdots$$

è un esempio di serie assolutamente convergente, poiché i termini della serie

$$\frac{\sin x}{1^2} + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{3^2} + \cdots$$

non sono maggiori dei termini corrispondenti della serie convergente

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots .$$

Un esempio di serie convergente, ma non assolutamente convergente, è il seguente

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

come il lettore potrà provare da solo.

Serie di funzioni; serie uniformemente convergenti. Nell'analisi abbiamo spesso a che fare con serie i cui termini sono funzioni di x. Nei paragrafi precedenti abbiamo già avuto esempi di questo genere, ad esempio la serie $1 + x + x^2 + x^3 + \cdots$. Per alcuni valori di x questa serie converge, ma per altri diverge. Particolarmente importanti nelle applicazioni sono serie di funzioni convergenti per tutti i valori di x appartenenti ad un certo intervallo, che può in particolare essere l'intero asse reale oppure la sua metà positiva e così via. Sorge allora la necessità di derivare tali serie termine per termine, integrarle, decidere se la loro somma è continua, e così via. Per il caso familiare della somma di un numero finito di termini esistono semplici regole generali. Sappiamo che la derivata di una somma di funzioni derivabili è uguale alla somma delle loro derivate, l'integrale di una somma di funzioni continue è la somma dei loro integrali e una somma di funzioni continue è essa stessa una funzione continua: tutto ciò vale per la somma di un numero finito di termini.

Ma per le serie infinite queste semplici regole in generale non sono più vere. Potremmo fornire molti esempi di serie convergenti di funzioni per le quali le regole di integrazione e derivazione termine per termine sono false. Allo stesso modo una serie di funzioni continue può risultare di somma discontinua. D'altra parte molte serie infinite si comportano come somme finite rispetto a queste regole.

Approfondite indagini su questa questione hanno dimostrato che queste regole possono ancora essere applicate se le serie infinite in questione non solo sono convergenti in ogni punto separato dell'intervallo di definizione (il dominio su cui varia x), ma se sono

uniformemente convergenti su tutto l'intervallo. In questo modo si cristallizzò nell'analisi matematica, a metà del XIX secolo, l'importante concetto della convergenza uniforme di una serie.

Consideriamo la serie

$$S(x) = u_0(x) + u_1(x) + u_2(x) + \cdots,$$

i cui termini sono funzioni definite sull'intervallo [a, b]. Supponiamo che per ogni valore separato di x nell'intervallo questa serie converga ad una certa somma $S(x)$. La somma dei primi n termini della serie

$$S_n(x) = u_0(x) + u_1(x) + \cdots + u_{n-1}(x)$$

è anche una certa funzione di x, definita su [a, b].

Introduciamo ora una grandezza η_n, che è uguale al estremo superiore dei valori[37] $| S(x) - S_n(x) |$, al variare di x nell'intervallo [a, b]. Questa grandezza è scritta come segue[38]

$$\eta_n = \sup_{a \leqslant x \leqslant b} | S_n(x) - S(x) |.$$

Nel caso in cui la grandezza $S(x) - S_n(x)$ raggiunga il suo valore massimo, cosa che si verificherà sicuramente ad esempio, quando $S(x)$ e $S_n(x)$ sono continue, allora η_n è semplicemente il massimo di $| S(x) - S_n(x) |$ su [a, b].

Dalla convergenza ipotizzata della nostra serie, abbiamo per ogni singolo valore di x nell'intervallo [a, b]

$$\lim_{n \to \infty} | S(x) - S_n(x) | = 0.$$

Ma la grandezza η_n potrebbe avvicinarsi allo zero oppure no. Se η_n, si avvicina a zero per $n \to \infty$, allora la serie si dice uniformemente convergente, e nel caso opposto non uniformemente convergente. Nello stesso senso è possibile parlare di convergenza uniforme o non uniforme di una successione di funzioni $S_n(x)$ senza necessariamente interpretarle come somme parziali di una serie.

Esempio 1. La serie di funzioni

$$\frac{1}{x+1} - \frac{1}{(x+1)(x+2)} - \frac{1}{(x+2)(x+3)} - \cdots,$$

che consideriamo definita solo per valori non negativi di x, cioè sulla semiretta [0, ∞), può essere scritta nella forma

$$\frac{1}{x+1} + \left(\frac{1}{x+2} - \frac{1}{x+1}\right) + \left(\frac{1}{x+3} - \frac{1}{x+2}\right) + \cdots,$$

da cui vediamo che le sue somme parziali sono uguali a

$$S_n(x) = \frac{1}{x+n}$$

E

$$\lim_{n\to\infty} S_n(x) = 0.$$

Pertanto la serie è convergente per tutti i x non negativi e ha la somma $S(x) = 0$. Inoltre,

$$\eta_n = \sup_{0\leqslant x<\infty} | S_n(x) - S(x) | = \sup_{0\leqslant x<\infty} \frac{1}{x+n} = \frac{1}{n} \to 0 \ (n \to \infty),$$

in modo che la serie sia uniformemente convergente a zero sul semiasse [0, ∞). La Figura 37 mostra i grafici di alcune somme parziali $S_n(x)$.

Esempio 2. La serie

$$x + x(x-1) + x^2(x-1) + \cdots$$

può essere scritta nella forma

$$x + (x^2 - x) + (x^3 - x^2) + \cdots,$$

da cui

$$S_n(x) = x^n,$$

e quindi

$$\lim_{n\to\infty} S_n(x) = \begin{cases} 0, \text{ if } 0 \leqslant x < 1; \\ 1, \text{ if } x = 1. \end{cases}$$

Quindi la somma della serie è discontinua sull'intervallo [0, 1] con una discontinuità nel punto $x = 1$. La quantità $| S_n(x) - S(x) |$ è inferiore all'unità per ogni x in [0, 1] ma per x vicino a $x = 1$ è arbitrariamente vicino all'unità. Quindi,

$$\eta_n = \sup_{0\leqslant x\leqslant 1} | S_n(x) - S(x) | = 1$$

per tutti $n = 1, 2, \cdots$. Quindi la serie converge in modo non uniforme sull'intervallo [0, 1]. La Figura 38 mostra alcuni grafici della funzione $S_n(x)$. Il grafico della somma della serie è costituito dal segmento $0 \leqslant x < 1$ dell'asse x omettendo il punto finale destro e dal punto (1,1).

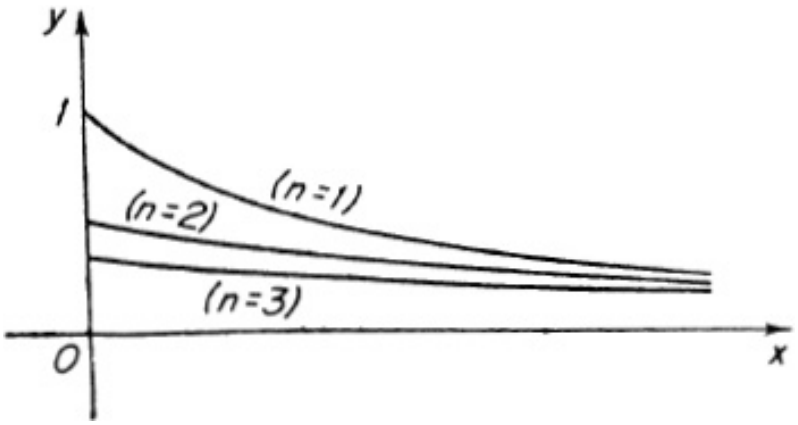


Figura 37.

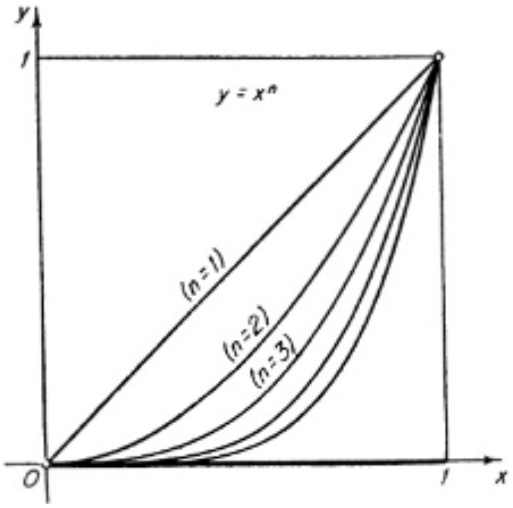


Figura 38.

Questo esempio mostra che la somma di una serie di funzioni continue convergenti in modo non uniforme può in effetti essere una funzione discontinua.

Se invece consideriamo la serie sull'intervallo $0 \leqslant x \leqslant q$ con $q < 1$, allora

$$\eta_n = \sup_{0 \leqslant x \leqslant q} | S_n(x) - S(x) | = \max_{0 \leqslant x \leqslant q} x^n = q^n \underset{n \to \infty}{\to} 0,$$

in modo che su questo intervallo la serie converge uniformemente e la sua somma è continua. Il fatto che la somma di una serie uniformemente convergente di funzioni continue sia essa stessa una funzione continua è una regola generale, come è stato sottolineato in precedenza, che può essere rigorosamente dimostrata.

Esempio 3. La somma dei primi n termini della serie $S_n(x)$ ha il grafico rappresentato dalla linea tratteggiata spessa in figura 39. Ovviamente per tutti gli n abbiamo $S_n(0) = 0$, ma se $0 < x \leqslant 1$, quindi per tutti gli $n \leqslant 1/x$, avremo $S_n(x) = 0$, e di conseguenza per x arbitrario nell'intervallo [0, 1],

$$S(x) = \lim_{n \to \infty} S_n(x) = 0.$$

D'altra parte,

$$\eta_n = \sup_{0 \leqslant x \leqslant 1} | S_n(x) - S(x) | = \sup | S_n(x) | = n^2.$$

Quindi la quantità η_n non si avvicina allo zero ma si avvicina addirittura all'infinito. Notiamo ora che la serie corrispondente a questa successione $S_n(x)$ non può essere integrata termine per termine sull'intervallo [0, 1], poiché

$$\int_0^1 S(x)\,dx = 0, \int_0^1 S_n(x)\,dx = \frac{1}{2} n^2 \frac{1}{n} = \frac{n}{2},$$

in modo che la serie

$$\int_0^1 S_1(x)\,dx + \int_0^1 [S_2(x) - S_1(x)]\,dx + \int_0^1 [S_3(x) - S_2(x)]\,dx + \cdots$$

si riduce alla serie divergente

$$\frac{1}{2} + \left(\frac{2}{2} - \frac{1}{2}\right) + \left(\frac{3}{2} - \frac{2}{2}\right) + \left(\frac{4}{2} - \frac{3}{2}\right) + \cdots.$$

Enunciamo senza dimostrazione le proprietà fondamentali delle serie uniformemente convergenti:

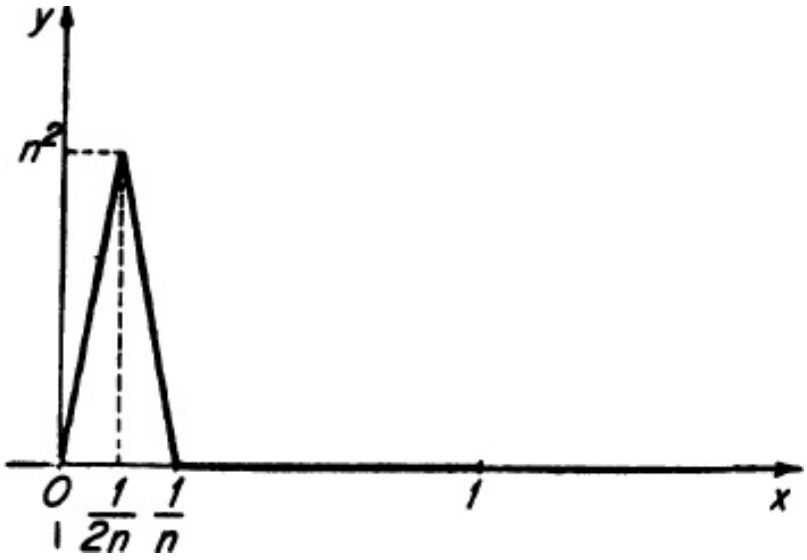


Figura 39.

1. La somma di una serie uniformemente convergente di funzioni continue sull'intervallo [*a*, *b*] è una funzione continua su questo intervallo.

2. Se la serie di funzioni continue

$$S(x) = u_0(x) + u_1(x) + u_2(x) + \cdots \tag{57}$$

converge uniformemente sull'intervallo [a, b], allora può essere integrata termine per termine su tale intervallo; cioè, per tutti i x_1, x_2 in [a, b] abbiamo l'uguaglianza

$$\int_{x_1}^{x_2} S(t)\,dt = \int_{x_1}^{x_2} u_0(t)\,dt + \int_{x_1}^{x_2} u_1(t)\,dt + \cdots.$$

3. Se sull'intervallo [a, b] la serie (57) converge e le funzioni $u_k(x)$ hanno derivate continue, allora l'uguaglianza

$$S'(x) = u_0'(x) + u_1'(x) + u_2'(x) + \cdots, \qquad (58)$$

ottenuta dalla derivazione termine per termine della (57) sarà valida sull'intervallo [a, b] se la serie a destra nella (58) converge uniformemente.

Serie di potenze. Nel §9, una funzione $f(x)$ definita su un intervallo [a, b] è stata detta analitica, se su tale intervallo ha derivate di ordine arbitrario e se in un intorno sufficientemente piccolo di un qualsiasi punto x_0 dell'intervallo [a, b] può essere sviluppata in una serie di Taylor convergente

$$f(x) = f(x_0) + \frac{f'(x_0)}{1}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots. \qquad (59)$$

Se introduciamo la notazione

$$a_n = \frac{f^{(n)}(x_0)}{n!},$$

questa serie può essere scritta nella forma seguente

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots. \qquad (60)$$

Una serie di questo tipo, dove i numeri a_1, a_2, ⋯ sono costanti indipendenti da x, è definita una serie di potenze.

Consideriamo come esempio la serie di potenze

$$1 + x + x^2 + x^3 + \cdots, \qquad (61)$$

i cui termini formano una progressione geometrica.

Sappiamo che per tutti i valori di x nell'intervallo $-1 < x < 1$ questa serie converge e la sua somma è pari a

$$S(x) = \frac{1}{1 - x}.$$

Per altri valori di x la serie diverge.

È anche facile vedere che la differenza tra la somma della serie e la somma dei suoi primi n termini è data dalla formula

$$S(x) - S_n(x) = \frac{x^n}{1 - x}, \qquad (62)$$

e se $-q \leqslant x \leqslant q$, dove q è un numero positivo inferiore all'unità, allora

$$\eta_n = \max | S(x) - S_n(x) | = \frac{q^n}{1 - q}.$$

Da ciò è chiaro che η_n si avvicina a zero all'aumentare di n in modo che la serie sia uniformemente convergente sull'intervallo $-q \leqslant x \leqslant q$, per tutti i valori positivi di $q < 1$.

È facile verificare che la funzione

$$S(x) = \frac{1}{1 - x}$$

ha una derivata dell'ordine n, che è uguale a

$$S^{(n)}(x) = \frac{n!}{(1 - x)^{n+1}},$$

da cui

$$S^{(n)}(0) = n!$$

e la somma dei primi n termini della serie di Taylor per la funzione $S(x)$ coincide esattamente per $x_0 = 0$ con la somma dei primi n termini della serie (59). Inoltre, sappiamo che il termine resto della formula, dato dall'uguaglianza (62), si avvicina a zero all'aumentare di n, per tutti i x sull'intervallo $-1 < x < 1$. Abbiamo quindi dimostrato che la serie (61) è la serie di Taylor della sua somma $S(x)$.

Notiamo un ulteriore fatto. Dall'intervallo di convergenza $-1 < x < 1$ della nostra serie, scegliamo un punto arbitrario x_0. È facile vedere che per tutti gli x sufficientemente vicini a x_0, cioè per tutti gli x che soddisfano la disuguaglianza

$$\frac{| x - x_0 |}{1 - x_0} < 1,$$

abbiamo l'uguaglianza

$$S(x) = \frac{1}{1-x} = \frac{1}{1-x_0}\frac{1}{\left(1-\frac{x-x_0}{1-x_0}\right)}$$
$$= \frac{1}{1-x_0}\left[1 + \frac{x-x_0}{1-x_0} + \left(\frac{x-x_0}{1-x_0}\right)^2 + \cdots\right]$$
$$= \frac{1}{1-x_0} + \frac{x-x_0}{(1-x_0)^2} + \frac{(x-x_0)^2}{(1-x_0)^3} + \cdots . \qquad (63)$$

Il lettore potrà provarlo senza difficoltà

$$\frac{S^{(n)}(x_0)}{n!} = \frac{1}{(1-x_0)^{n+1}} .$$

Di conseguenza, la serie (63) è la serie di Taylor della sua somma $S(x)$ e converge ad essa in un intorno sufficientemente piccolo di un qualsiasi punto x_0 appartenente all'intervallo di convergenza della (61). Poiché il punto x_0 è arbitrario, ciò significa che la funzione $S(x)$ è analitica sull'intervallo.

Tutti questi fatti che abbiamo osservato per la particolare serie di potenze (61) valgono per serie di potenze arbitrarie.[39] Vale a dire, per ogni serie di potenze della forma (60) dove le costanti a_k sono scelte da una determinata legge, esiste un certo numero non negativo R (che può anche essere infinito), chiamato raggio di convergenza della serie (60), con le seguenti proprietà:

1. Per tutti i valori di x dall'intervallo $x_0 - R < x < x_0 + R$, che è chiamato intervallo di convergenza, la serie converge e la sua somma $S(x)$ è una funzione analitica di x nel suo intervallo. Qui la convergenza è uniforme per ogni intervallo $[a, b]$ che giace completamente all'interno dell'intervallo di convergenza. La serie stessa è la serie di Taylor della sua somma.

2. Ai punti finali dell'intervallo di convergenza, la serie può convergere o divergere, a seconda del suo carattere individuale. Ma certamente divergerà al di fuori dell'intervallo chiuso $x_0 - R \leqslant x \leqslant x_0 + R$.

Suggeriamo al lettore di considerare le serie di potenze

$$1 + \frac{x}{1} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots ,$$
$$1 + x + 2!x^2 + 3!x^3 + \cdots ,$$
$$1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots$$

e convincersi che i loro raggi di convergenza siano rispettivamente infinito, zero e unità.

Per la definizione data in precedenza ogni funzione analitica può essere estesa, in un intorno sufficientemente piccolo di un punto arbitrario in cui è definita, in una serie di potenze che

converge alla funzione. Viceversa, da quanto detto segue che la somma di ogni serie di potenze il cui raggio di convergenza non è zero è una funzione analitica nel suo intervallo di convergenza.

Quindi vediamo che le serie di potenze sono organicamente connesse con le funzioni analitiche. Potremmo anche dire che nel loro intervallo di convergenza le serie di potenze sono il mezzo naturale per rappresentare le funzioni analitiche, e di conseguenza sono anche il mezzo naturale per approssimare le funzioni analitiche mediante polinomi algebrici.[40]

Ad esempio dal fatto che la funzione $1/(1 - x)$ può essere ampliata nella serie di potenze

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots,$$

che converge sull'intervallo $-1 < x < 1$, ne consegue che la serie di potenze è uniformemente convergente su un intervallo arbitrario $-a \leqslant x \leqslant a$ con $a < 1$, e ciò implica la possibilità di approssimare la funzione sull'intero intervallo $[-a, a]$ mediante le somme parziali della serie con qualsiasi preassegnata grado di precisione.

Supponiamo di dover approssimare la funzione $1/(1 - x)$ mediante polinomi sull'intervallo $[-\frac{1}{2}, \frac{1}{2}]$ con una precisione di 0,01. Notiamo che per tutti gli x in questo intervallo abbiamo la disuguaglianza

$$\left| \frac{1}{1-x} - 1 - x - \cdots - x^n \right| = |\, x^{n+1} + x^{n+2} + \cdots |$$
$$\leqslant |\, x \,|^{n+1} + |\, x \,|^{n+2} + \cdots \leqslant \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots = \frac{1}{2^n},$$

e poiché $2^6 = 64$, e $2^7 = 128$, il polinomio desiderato, approssimando la funzione su tutto l'intervallo $[-\frac{1}{2}, \frac{1}{2}]$ con una precisione di 0,01, avrà la forma

$$\frac{1}{1-x} \approx 1 + x + x^2 + \cdots + x^7.$$

Notiamo un'altra proprietà estremamente preziosa delle serie di potenze: esse possono essere differenziate a termine ovunque nell'intervallo di convergenza. Questa proprietà trova un'applicazione estremamente ampia nella soluzione di vari problemi di matematica.

Ad esempio, sia necessario trovare la soluzione dell'equazione differenziale $y' = y$ nella condizione ausiliaria $y(0) = 1$. Cercheremo la soluzione sotto forma di una serie di potenze,

$$y = a_0 + a_1 x + a_2 x^2 + \cdots.$$

A causa della condizione ausiliaria, dobbiamo impostare $a_0 = 1$. Supponendo che questa serie converga, possiamo differenziarla a termine; di conseguenza otteniamo

$$y' = a_1 + 2a_2 x + 3a_3 x^2 + \cdots.$$

Se sostituiamo queste due serie nell'equazione differenziale e uguagliamo i coefficienti per ciascuna delle potenze di x, otteniamo

$$a_k = \frac{1}{k!} \quad (k = 1, 2, \cdots)$$

e la soluzione desiderata ha la forma

$$y = 1 + \frac{x}{1} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots .$$

È noto che questa serie converge per tutti i valori di x e che la sua somma è pari a $y = e^x$.

In questo caso abbiamo ottenuto una serie la cui somma è una funzione elementare ben nota. Ma questo non sempre accade; può risultare che una serie di potenze convergente così ottenuta abbia una somma che non è una funzione elementare. Un esempio è la serie

$$y_p(x) = x^p \left[1 - \frac{x^2}{2(2p+2)} + \frac{x^4}{2 \cdot 4(2p+2)(2p+4)} - \cdots\right],$$

ottenuto come soluzione dell'equazione differenziale di Bessel, che è di grande importanza nelle applicazioni. In questo modo le serie di potenze possono servire a definire funzioni.

Letture consigliate

R. Courant, Differential and integral calculus, 2 vols., Interscience, New York, 1938.

A. Dresden, Introduction to the calculus, Henry Holt, New York, 1940.

F. Klein, Elementary mathematics from an advanced standpoint. Arithmetic-algebra-analysis, Dover, New York, 1953.

K. Knopp, Infinite sequences and series, Dover, New York, 1956.

Note al capitolo II

[1] Questa sezione è basata sul bellissimo saggio dell'accademico S. I. Vavilov “Galileo” (Grande Enciclopedia Sovietica, Volume 10, 1952).

[2] Oggi la formula (1) si deduce dalle leggi generali della meccanica, ma storicamente fu proprio questa formula che, dopo essere stata stabilita sperimentalmente da Galileo, servì come parte dell'accumulazione di esperienza successivamente generalizzata da quelle leggi.

[3] Il numero x è l'ascissa e y è l'ordinata del punto M.

[4] Ciò non significa che tale espressione sia impossibile. Nel Capitolo XII mostreremo come ottenere un'unica formula.

[5] Naturalmente, l'affermazione di cui sopra sarà completamente chiara al lettore solo dopo aver dato una definizione precisa di cosa si intende esattamente in matematica con il termine "formula" ed "espressione analitica".

[6] Se nelle uguaglianze ovvie $(k + 1)^3 - k^3 = 3k^2 + 3k + 1$, per i diversi valori $k = 1, 2, \cdots, n - 1$, sommiamo separatamente i lati sinistro e destro, otteniamo l'equazione

$$n^3 - 1 = 3\sigma_n + \frac{3(n-1)n}{2} + n - 1$$

dove $\sigma_n = 1^2 + 2^2 + \ldots + (n-1)^2$. Risolvendo questa equazione per σ_n, otteniamo

$$\sigma_n = \frac{(n-1)n(2n-1)}{6}.$$

[7] Ad esempio, se $\epsilon = 0{,}001$, possiamo prendere $N = 500$. Infatti, poiché

$$\frac{1}{6n^2} < \frac{1}{2n}$$

per ogni intero positivo n, quindi

$$|\alpha_n| = \left|\frac{1}{6n^2} - \frac{1}{2n}\right| = \frac{1}{2n} - \frac{1}{6n^2} < \frac{1}{2n} < 0.001$$

per n arbitrario > 500. Allo stesso modo sarebbe possibile assegnare valori arbitrariamente piccoli ϵ, ad esempio:

$$\epsilon_1 = 0.0001, \qquad \epsilon_2 = 0.00001, \cdots,$$

e per ciascuno di essi scegliere, come sopra, i valori opportuni $N = N_1, N_2, \cdots$.

[8] A questo proposito va osservato che l'addizione, la sottrazione, la moltiplicazione e la divisione (esclusa la divisione per zero) di numeri razionali, cioè numeri della forma p/q dove p e q sono numeri interi, porta a numeri razionali. Ma questo non è necessariamente il caso dell’operazione di presa di un limite. Il limite di una successione di numeri razionali può essere un numero irrazionale.

[9] Non è importante che il poligono sia regolare. L'unica caratteristica essenziale è che il lato maggiore del poligono inscritto variabile si avvicini allo zero.

[10] Basta porre la funzione uguale all'unità nei punti irrazionali.

[11] Resta inteso che qui si parla del caso in cui il limite in questione esiste effettivamente. Se questo limite non esiste allora diciamo che nel punto x la funzione non ha derivata.

[12] Qui assumiamo sempre che $h \neq 0$.

[13] Qui Δx non è mai uguale a zero.

[14] L'addendo finale qui si avvicina a zero per $\Delta x \to 0$, poiché $\Delta v/\Delta x$ si avvicina a un numero finito, uguale alla derivata v', che si supponeva esistesse fin dall'inizio, e $\Delta u \to 0$, poiché la funzione u, che si suppone derivabile, è continua.

[15] Nel dedurre questa formula abbiamo tacitamente assunto che, man mano che Δx si avvicina a zero, Δu non è mai uguale a zero. Ma la formula rimane vera anche quando questa ipotesi non regge.

[16] In senso stretto, la "convessità verso l'alto" è quella proprietà della curva che consiste nel trovarsi al di sopra (più precisamente "non al di sotto") della corda che congiunge due suoi punti qualsiasi; analogamente, per la "convessità verso il basso" (detta anche semplicemente "concavità"), la curva non giace al di sopra delle sue corde.

[17] Nei casi più complicati, dove la stessa derivata seconda cambia segno, il problema di determinare il carattere del punto stazionario si risolve mediante la formula di Taylor (§9).

[18] Questa formula fornisce tutte le soluzioni.

[19] Naturalmente questi argomenti danno solo un'interpretazione geometrica del teorema e non costituiscono in alcun modo una dimostrazione rigorosa.

[20] La formula (23) può essere derivata mediante una semplice applicazione del teorema del valore medio alla funzione

$$f(x) = \phi(x) - \frac{\phi(b) - \phi(a)}{\psi(b) - \psi(a)}\psi(x).$$

[21] Con i simboli $[a, b]$ e (a, b) indichiamo gli insiemi di valori di x che soddisfano le disuguaglianze $a \leqslant x \leqslant b$ e $a < x < b$ rispettivamente.

[22] La stessa regola è valida per trovare il limite di un'espressione frazionaria in cui il numeratore e il denominatore si avvicinano entrambi all'infinito. Questo metodo, molto comodo per trovare tali limiti (o, come diciamo, per rimuovere le indeterminazioni), sarà utilizzato, ad esempio, nel §3 del capitolo XII.

[23] Questa è solo una delle forme possibili per il termine residuo $R_{n+1}(x)$.

[24] Per il teorema del valore medio

$$\phi(t) - \phi(t_0) = \phi'(v)(t - t_0) = 0,$$

quando v si trova tra t e t_0. Pertanto $\phi(t) = \phi(t_0) = \text{const}$ per tutti gli t.

[25] È possibile dimostrare matematicamente, senza ricorrere ad esempi tratti dalla meccanica, che se la funzione $f(x)$ è continua (e anche se è discontinua ma integrabile secondo Lebesgue; vedi Capitolo XV) sull'intervallo $[a, b]$, allora esiste una primitiva $F(x)$ che soddisfa l'uguaglianza (30).

[26] Questa formula è stata generalizzata in vari modi (vedi ad esempio §13, la formula di Ostrogradskij).

[27] Le scoperte di Newton e Leibniz furono fatte indipendentemente.

[28] Ad esempio, da $a = b$ e $b = c$ consegue formalmente che $a = c$. Ma in pratica questa relazione si presenta come segue: Dal fatto che $a = b$ è noto con precisione fino a ε e $b = c$ è noto con la stessa precisione, ne consegue che $a = c$ è noto con precisione fino a 2ε.

[28a] Per $x > 0$, $(\ln |x|)' = (\ln x)' = 1/x$; per $x < 0$, $(\ln |x|)' = [\ln(-x)]' = [1/(-x)](-1) = 1/x$.

[29] In generale una funzione $Ax + By + C$, dove A, B, C sono costanti, è chiamata funzione lineare di x e y. Se $C = 0$ si dice funzione lineare omogenea. Qui omettiamo la parola "omogeneo".

[30] Assumiamo che $F(x, y)$ abbia derivate continue rispetto a x e y.

[31] Si tratta ovviamente di un massimo raggiunto nel dominio di definizione della funzione $F(x, y)$. Si presuppone che le funzioni $f(x, y)$ e $\phi(x, y)$ siano differenziabili.

[32] Nel corso di matematica superiore di V. I. Smirnov il lettore troverà un semplice esempio in cui questa particolare caratteristica della situazione porterebbe alla perdita della soluzione se applichiamo meccanicamente il metodo di Lagrange e non consideriamo, oltre ai punti sopra menzionati, un punto in cui vale non solo la (45) ma anche:

$$\phi'_x(x_0, y_0) = 0, \qquad \phi'_y(x_0, y_0) = 0$$

[33] Il diametro di una regione è definito come il estremo superiore della distanza tra due punti della regione.

[34] Più precisamente attraverso una superficie cilindrica avente per base il contorno e con altezza pari all'unità.

[35] Poiché per piccoli spostamenti lungo la curva il differenziale della coordinata y è pari a $\cos(\bar{n}, \bar{x})$ ds e il differenziale dx è pari a $-\cos(\bar{n}, \bar{y})$ ds, quest'ultimo integrale viene spesso scritto nella forma

$$\int_{\Gamma} [P(x, y)\, dy - Q(x, y)\, dx].$$

[36] Notiamo che è anche possibile dare definizioni generalizzate di somma di una serie, in virtù delle quali è possibile assegnare a certe serie divergenti un concetto più o meno naturale di "somma generalizzata". Tali serie si dicono sommabili. Talvolta sono utili le operazioni con somme generalizzate di serie divergenti.

[37] Vedi Capitolo XV.

[38] sup è l'abbreviazione della parola latina supremum (il più alto).

[39] Per informazioni più dettagliate su questo punto vedere il Capitolo IX.

[40] Approssimazioni che vanno oltre i limiti dell'intervallo di convergenza di una serie di potenze richiedono altri metodi. (Vedi capitolo XII.)

Nota editoriale

[E1] Nel §7 è corretto un refuso dell'edizione di partenza: la funzione $\exp(-x^2)$ ha il massimo in $x = 0$, non in $x = 1$. Il rinvio al grafico della funzione lineare indica la figura 5 del §2 di questo capitolo. Nel §12, nel problema di massimo vincolato, si cerca il massimo di $f(x, y)$ sulla curva Γ: il testo di partenza riportava invece «di x e y». Sono inoltre corretti errori di trascrizione dell'edizione elettronica, fra cui il valore 0.05 nella tabella 1, verificati sulla copia DjVu.

CAPITOLO III

Geometria analitica

B. N. Delone

§ 1. Introduzione

Nella prima metà del XVII secolo sorse una branca completamente nuova della matematica, la cosiddetta geometria analitica, che stabiliva una connessione tra le curve in un piano e le equazioni algebriche in due incognite.

In questo modo accadde in matematica un evento piuttosto raro: in uno o due decenni apparve un grande ramo completamente nuovo della matematica basato su un concetto molto semplice, che fino ad allora non aveva ricevuto la dovuta attenzione. La comparsa della geometria analitica nella prima metà del XVII secolo non fu casuale. Il passaggio in Europa ai nuovi metodi di produzione capitalistici richiese il progresso di tutta una serie di scienze. Poco tempo prima, Galileo e altri scienziati avevano creato la meccanica contemporanea, in tutti i campi delle scienze naturali si accumulavano dati sperimentali, si perfezionavano i mezzi di osservazione e al posto delle obsolete teorie scolastiche se ne creavano di nuove. In astronomia, tra gli scienziati più eminenti avevano finalmente trionfato gli insegnamenti di Copernico. Il rapido sviluppo della navigazione a lungo raggio richiedeva con insistenza la conoscenza dell'astronomia e degli elementi di meccanica.

L'arte della guerra richiedeva anche la meccanica. Ellissi e parabole, le cui proprietà geometriche come sezioni coniche erano già ben note agli antichi greci quasi 2000 anni prima, cessarono di essere solo una parte della geometria, come lo erano per i greci. Dopo che Keplero ebbe scoperto che i pianeti girano attorno al sole secondo ellissi, e Galileo che un sasso lanciato in aria traccia una parabola, fu necessario calcolare queste ellissi e trovare le parabole lungo le quali volano le pallottole di un cannone; era necessario scoprire la legge per cui la pressione atmosferica, scoperta da Pascal, diminuisce con l'altezza; era infatti necessario calcolare i volumi dei vari corpi, e così via.

Tutte queste domande diedero vita quasi simultaneamente a tre scienze matematiche completamente nuove: la geometria analitica, il calcolo differenziale e il calcolo integrale, compresa la soluzione delle equazioni differenziali più semplici.

Questi tre nuovi campi cambiarono qualitativamente il volto dell’intera matematica. Permisero di risolvere problemi mai nemmeno immaginati prima.

Nella prima metà del XVII secolo, cioè all’inizio del 1600, un gruppo dei matematici più eminenti era già vicino all’idea della geometria analitica, ma furono due, in particolare, a comprendere chiaramente la possibilità di creare una nuova branca della matematica. Questi erano Pierre Fermat, consigliere del parlamento della città francese di Tolosa e matematico di

fama mondiale, e il famoso filosofo francese René Descartes. A Cartesio viene attribuito il merito di essere il principale creatore della geometria analitica. Fu lui che, come filosofo, pose la questione in tutta la sua generalità. Cartesio pubblicò il grande trattato filosofico “Discorso sul metodo di condurre rettamente la ragione e di ricercare la verità nelle scienze, con applicazioni: diottrica, meteorologia e geometria”.

L'ultima parte di quest'opera, intitolata “Geometria” e pubblicata nel 1637, contiene una presentazione sufficientemente completa, anche se alquanto confusa, della teoria matematica che da allora è stata chiamata geometria analitica.

§ 2. Le due idee fondamentali di Cartesio

Cartesio desiderava creare un metodo che potesse essere applicato ugualmente bene alla soluzione di tutti i problemi di geometria, cioè che fornisse un metodo generale per la loro soluzione. La teoria di Cartesio si basa su due concetti: il concetto di coordinate e il concetto di rappresentare mediante il metodo delle coordinate qualsiasi equazione algebrica con due incognite sotto forma di una curva nel piano.

Il concetto di coordinate. Per coordinate di un punto nel piano Cartesio intende l'ascissa e l'ordinata di questo punto, cioè i valori numerici x e y delle sue distanze (con segni corrispondenti) da due rette reciprocamente perpendicolari (assi coordinati) scelte in questo piano (vedi Capitolo II). Il punto di intersezione degli assi delle coordinate, cioè il punto avente coordinate (0, 0) è chiamato origine.

Con l'introduzione delle coordinate Cartesio costruì, per così dire, una “aritmetizzazione” del piano. Invece di determinare geometricamente un punto qualsiasi, è sufficiente fornire una coppia di numeri x, y e viceversa (figura 1).

La corrispondenza fra equazioni in due incognite e curve del piano. Il secondo concetto di Cartesio è il seguente. Fino ai tempi di Cartesio, dove veniva data un'equazione algebrica in due incognite $F(x, y) = 0$, si diceva che il problema era indeterminato, poiché dall'equazione era impossibile determinare queste incognite; ad uno di essi si potrebbe assegnare qualsiasi valore, ad esempio a x, e sostituirlo nell'equazione; il risultato fu un'equazione con una sola incognita y, per la quale, in generale, l'equazione poteva essere risolta. Allora questo x scelto arbitrariamente insieme all'y così ottenuto soddisferebbero l'equazione data. Di conseguenza, un’equazione così “indeterminata” non è stata considerata interessante.

Cartesio vedeva la questione in modo diverso. Propose di considerare in un'equazione a due incognite x come l'ascissa di un punto e il corrispondente y come la sua ordinata. Se poi variamo l'incognita x, ad ogni valore di x viene calcolato il corrispondente y dall'equazione, così da ottenere, in generale, un insieme di punti che formano una curva (figura 2).[1]

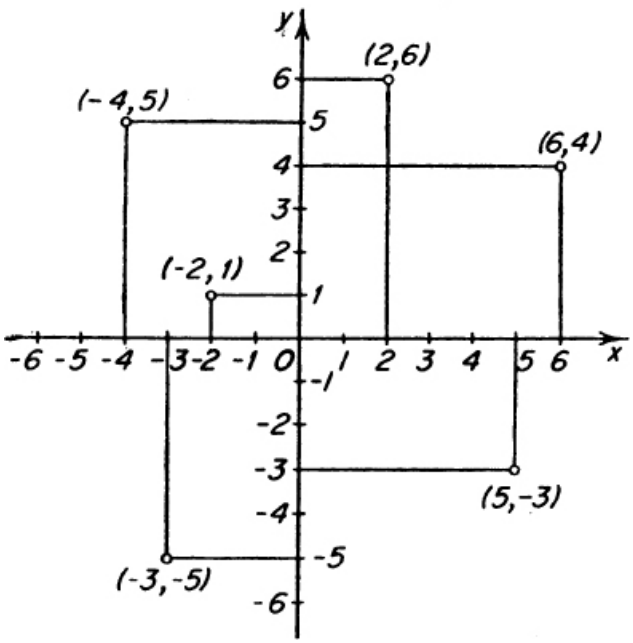


Figura 1.

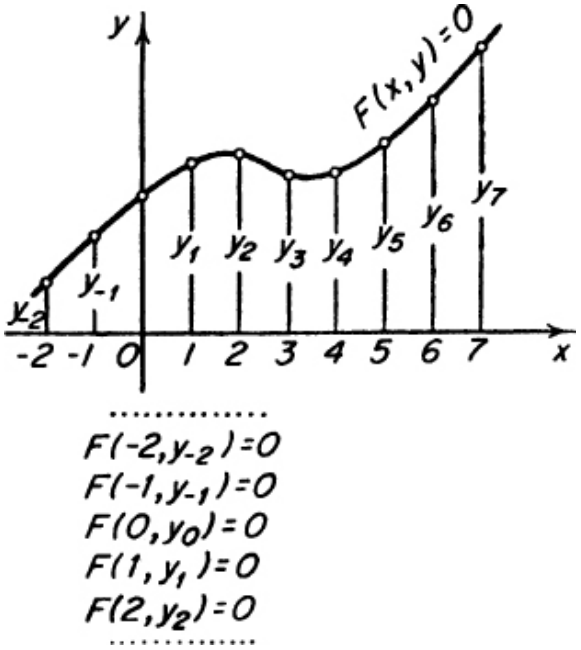


Figura 2.

Pertanto, ad ogni equazione algebrica a due variabili, $F(x, y) = 0$, corrisponde una curva del piano completamente determinata, cioè una curva che rappresenta la totalità di tutti quei punti del piano le cui coordinate soddisfano l'equazione $F(x, y) = 0$.

Questa osservazione di Cartesio ha aperto una scienza completamente nuova.

I problemi fondamentali risolti dalla geometria analitica e la definizione di geometria analitica. La geometria analitica offre la possibilità: (1) di risolvere problemi di costruzione mediante calcolo (vedi ad esempio la divisione di un segmento in un dato rapporto, vedi §3); (2) di trovare l'equazione delle curve definite da una proprietà geometrica (ad esempio, di

un'ellisse definita dalla condizione che la somma delle distanze da due punti dati sia costante, vedi §7); (3) di dimostrare algebricamente nuovi teoremi geometrici (vedi, ad esempio, la derivazione della teoria dei diametri di Newton, §6); (4) viceversa, di rappresentare geometricamente un'equazione algebrica, per chiarirne le proprietà algebriche (vedi, ad esempio, la soluzione delle equazioni di terzo e quarto grado dall'intersezione di una parabola con un cerchio, §5).

La geometria analitica è quindi quella parte della matematica che, applicando il metodo delle coordinate, indaga gli oggetti geometrici con mezzi algebrici.

§ 3. Problemi elementari

Le coordinate di un punto che divide un segmento in un dato rapporto. Date le coordinate (x_1, y_1) e (x_2, y_2) di due punti M_1 e M_2, troviamo le coordinate (x, y) del punto M che divide il segmento M_1M_2 nel rapporto m a n (figura 3). Dalla similitudine dei triangoli ombreggiati otteniamo:

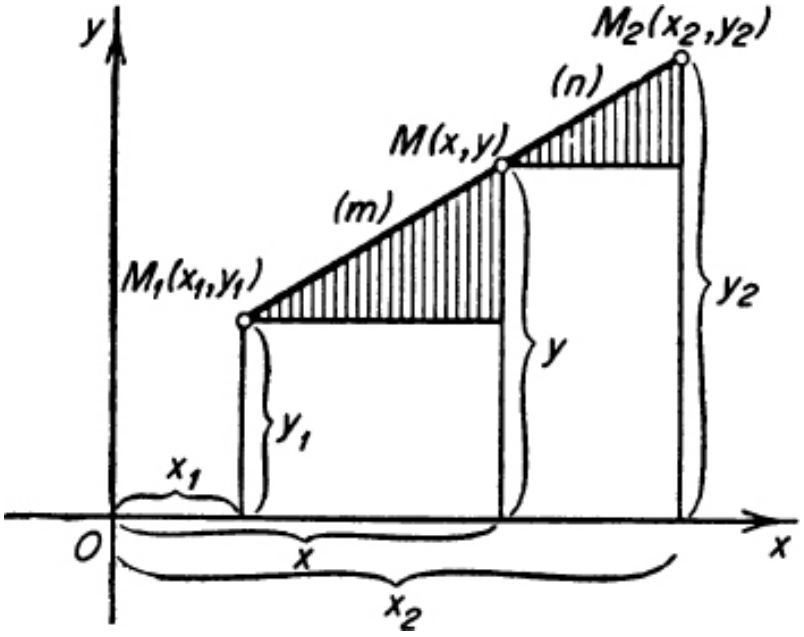


Figura 3.

$\frac{x - x_1}{x_2 - x} = \frac{m}{n}$, da cui $x = \frac{nx_1 + mx_2}{m + n}$, $\frac{y - y_1}{y_2 - y} = \frac{m}{n}$, da cui $y = \frac{ny_1 + my_2}{m + n}$

Distanza tra due punti. Troviamo la distanza tra i punti M_1 e M_2, le cui coordinate sono (x_1, y_1) e (x_2, y_2) rispettivamente. Dal triangolo rettangolo ombreggiato (figura 4), si ottiene dal teorema di Pitagora

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

L'area di un triangolo. Troviamo l'area S del triangolo $M_1M_2M_3$ (figura 5) se le coordinate dei suoi vertici sono rispettivamente (x_1, y_1), (x_2, y_2), (x_3, y_3). Considerando l'area del triangolo

come la somma delle aree dei trapezi con basi y_1, y_3 e y_3, y_2 meno l'area del trapezio con basi y_1, y_2 e scrivendo il prodotto $-(y_1 + y_2)(x_2 - x_1)$ nella forma $(y_1 + y_2)(x_1 - x_2)$, otteniamo

$$S = \tfrac{1}{2}[(y_1 + y_2)(x_1 - x_2) + (y_2 + y_3)(x_2 - x_3) + (y_3 + y_1)(x_3 - x_1)].$$

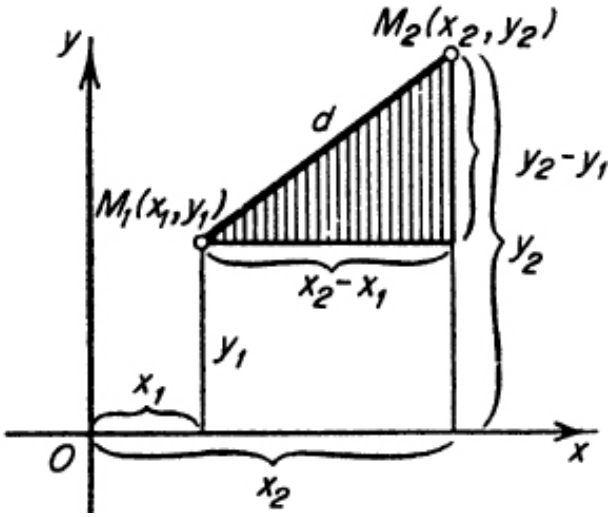


Figura 4.

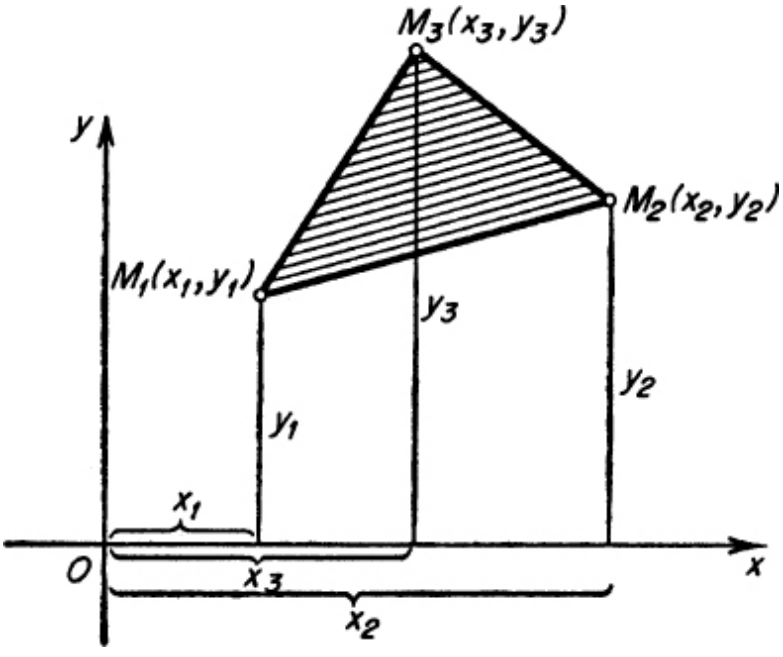


Figura 5.

In questi problemi non resta che verificare che le formule derivate rimangano valide senza alcuna variazione nei casi in cui una o più coordinate o le loro differenze sono negative. Tale verifica è facilmente conseguibile.

Determinazione dei punti di intersezione di due curve. Basandosi sulla seconda idea fondamentale che l'equazione $F(x, y) = 0$ rappresenta una curva, è particolarmente semplice trovare i punti di intersezione di due curve. Per trovare le coordinate dei punti di intersezione di due curve è ovviamente necessario risolvere contemporaneamente le equazioni che le rappresentano. La coppia di numeri x, y ottenuta dalla soluzione simultanea di queste due equazioni determinerà il punto le cui coordinate soddisfano entrambe le equazioni, cioè il punto che giace sia sulla prima che sulla seconda curva, e questo è il punto della loro intersezione.

La soluzione dei problemi geometrici con gli strumenti della geometria analitica, come vediamo, è molto comoda per scopi pratici, soprattutto perché ogni soluzione viene subito ottenuta nella comoda forma numerica. Una tale geometria, una tale scienza, era esattamente ciò che mancava a quel tempo.

§ 4. Studio delle curve rappresentate da equazioni di primo e secondo grado

Equazione di primo grado. Facendo uso della sua seconda idea, Cartesio esaminò innanzitutto quali curve corrispondono a un'equazione di primo grado,

$$Ax + By + C = 0, \qquad (1)$$

cioè, a un'equazione in cui A, B, C sono coefficienti numerici con A e B non entrambi zero. Si è scoperto che nel piano una retta corrisponde sempre a tale equazione.

Dimostreremo che l'equazione (1) rappresenta sempre una retta e, viceversa, che ad ogni retta del piano corrisponde un'equazione completamente determinata della forma (1). Supponiamo infatti, ad esempio, che $B \neq 0$; quindi l'equazione (1) può essere risolta per y

$$y = kx + l,$$

dove $k = -\frac{A}{B}; l = -\frac{C}{B}$.

Esaminiamo innanzitutto l'equazione $y = kx$. Rappresenta ovviamente una retta passante per l'origine e formante un angolo ϕ con l'asse x la cui tangente $\tan \phi$ è k (figura 6). In effetti, l'equazione può essere

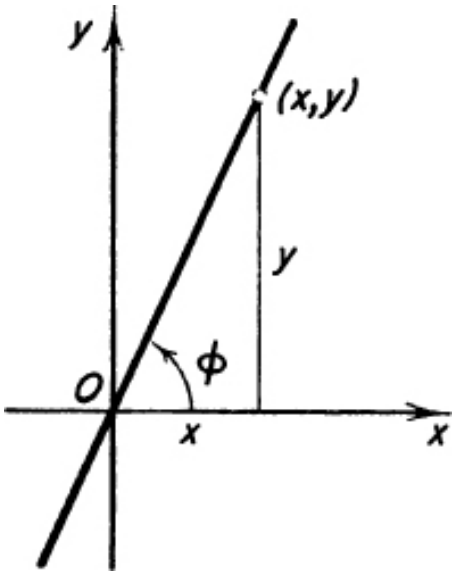


Figura 6.

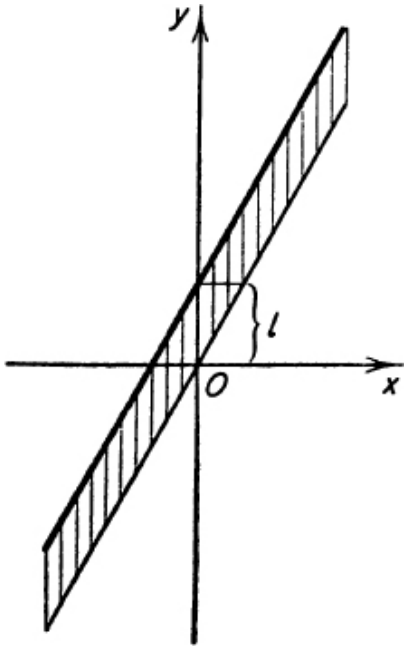


Figura 7.

scritto come $y/x = k$, in modo che le coordinate di ogni punto (x, y) sulla retta soddisfano l'equazione, e le coordinate di nessun punto $(x, \bar{y})$ che non giace sulla retta soddisfano l'equazione, poiché per tale punto y/x sarà maggiore o minore di k. Inoltre, se tan $\phi > 0$, allora per questa linea sia x che y sono positivi o entrambi negativi e se tan $\phi < 0$ i loro segni sono opposti.

Pertanto, l'equazione $y = kx$ rappresenta una retta passante per l'origine O, e di conseguenza anche l'equazione $y = kx + l$ rappresenta una retta, cioè quella che si ottiene dalla retta precedente mediante la traslazione parallela tale per cui l'ordinata di ciascuno dei suoi punti è aumentata di l (figura 7).

Le formule precedentemente derivate delle coordinate del punto che divide un segmento in un dato rapporto, della distanza tra due punti dati e dell'area di un triangolo, nonché le informazioni sull'equazione di una retta ci permettono già di risolvere un gran numero di problemi.

Equazione della retta passante per uno o due punti dati. Sia M_1 il punto di coordinate x_1, y_1 e sia k un numero dato. L'equazione $y = kx + l$ rappresenta una retta che forma con l'asse Ox un angolo la cui tangente è uguale a k e che interseca l'asse Oy a distanza l da O. Scegliamo l in modo tale che questa linea passi per il punto (x_1, y_1). Per questo, le coordinate del punto M_1, devono soddisfare l'equazione, cioè dobbiamo avere $y_1 = kx_1 + l$, da cui $l = y_1 - kx_1$.

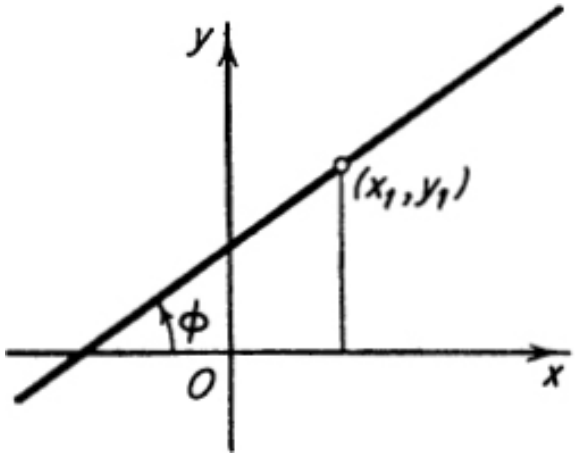


Figura 8.

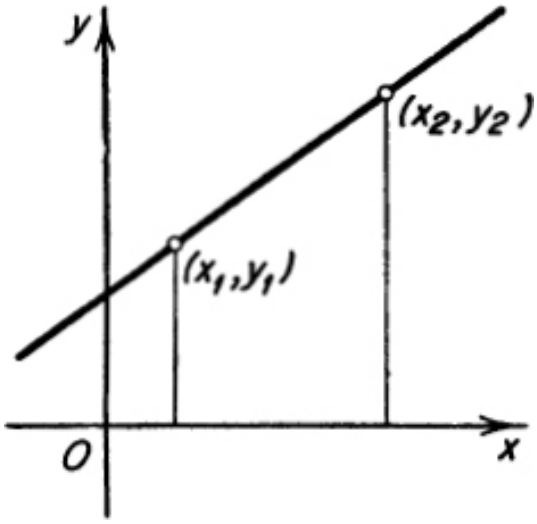


Figura 9.

Sostituendo questo valore a l, otteniamo l'equazione della retta che passa per il punto dato (x_1, y_1) e forma con l'asse Ox un angolo la cui tangente è uguale a k (figura 8). Questa equazione è $y = kx + y_1 - kx_1$ o

$$y - y_1 = k(x - x_1).$$

Esempio. Sia l'angolo tra la retta e l'asse Ox pari a 45°, e il punto M abbia coordinate (3,7); allora l'equazione della retta corrispondente (poiché tan 45° = 1) sarà: $y - 7 = 1 \cdot (x - 3)$ oppure $x - y + 4 = 0$.

Se richiediamo che la retta passante per il punto (x_1, y_1) passi anche per il punto (x_2, y_2), ne consegue che la condizione $y_2 - y_1 = k(x_2 - x_1)$ deve essere imposta su k. Trovando da questa k e sostituendolo nell'equazione precedente, otteniamo l'equazione della retta passante per due punti dati (figura 9):

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}.$$

Risultato di Cartesio relativo alle equazioni di secondo grado. Cartesio investigò anche la questione di quali tipi di curve nel piano siano rappresentate dall'equazione di secondo grado con due variabili la cui forma generale è

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0,$$

e dimostrò che tale equazione, in generale, rappresenta un'ellisse, un'iperbole o una parabola; cioè curve ben note ai matematici dell'antichità.

Questi sono i risultati più importanti di Cartesio. Tuttavia, il suo libro era lungi dall'essere limitato a questi argomenti; studiò anche le equazioni di numerosi luoghi geometrici interessanti, esaminò alcuni teoremi sulla trasformazione delle equazioni algebriche, menzionò senza prove la sua famosa legge dei segni per il numero di radici positive di un'equazione le cui radici sono tutte reali (vedi Capitolo IV, §4) e, infine, presentò un notevole metodo per determinare le radici reali delle equazioni di terzo e quarto grado dall'intersezione della parabola $y = x^2$ con cerchi.

§ 5. Il metodo di Cartesio per risolvere le equazioni algebriche di terzo e quarto grado

Trasformazione di equazioni di terzo e quarto grado in un'equazione di quarto grado che non coinvolga il termine x^3. Mostreremo che la soluzione di un'equazione arbitraria di terzo o quarto grado può essere ridotta alla soluzione di un'equazione della forma

$$x^4 + px^2 + qx + r = 0. \qquad (2)$$

Sia $z^3 + az^2 + bz + c = 0$. Sostituendo $z = x - a/3$, otteniamo

$$(x - a/3)^3 + a(x - a/3)^2 + b(x - a/3) + c = 0.$$

I termini x^2 nell'espansione delle parentesi si annulleranno, in modo da ottenere un'equazione della forma $x^3 + px + q = 0$. Moltiplicando questa equazione per x la portiamo alla forma (2) con $r = 0$, che ammette anche una radice $x_4 = 0$.

Un'equazione di quarto grado $z^4 + az^3 + bz^2 + cz + d = 0$ può essere ridotta alla forma (2) mediante la sostituzione $z = x - a/4$. Pertanto, la soluzione di tutte le equazioni di terzo e quarto grado può essere ridotta alla soluzione di un'equazione della forma (2).

La soluzione delle equazioni di terzo e quarto grado mediante l'intersezione di un cerchio con la parabola $y = x^2$. Deriviamo prima l'equazione di una circonferenza con centro (a, b) e raggio R. Se (x, y) è uno qualsiasi dei suoi punti, allora il quadrato della sua distanza dal punto (a, b) è uguale a $(x - a)^2 + (y - b)^2$ (vedi §3). Pertanto, l'equazione del cerchio in questione è

$$(x - a)^2 + (y - b)^2 = R^2.$$

Ora proviamo a trovare i punti di intersezione di questa circonferenza con la parabola $y = x^2$. Per fare ciò, in virtù di quanto detto nel §3, è necessario risolvere contemporaneamente l'equazione di tale circonferenza

$$x^2 + y^2 - 2ax - 2by + a^2 + b^2 - R^2 = 0$$

e l'equazione della parabola

$$y = x^2.$$

Sostituendo y dalla seconda equazione nella prima, otteniamo un'equazione di quarto grado in x:

$$x^2 + x^4 - 2ax - 2bx^2 + a^2 + b^2 - R^2 = 0$$

O

$$x^4 + (1 - 2b)\, x^2 - 2ax + a^2 + b^2 - R^2 = 0.$$

Se scegliamo a, b e R^2 in modo tale che

$$1 - 2b = p, - 2a = q, a^2 + b^2 - R^2 = r,$$

allora si ottiene esattamente l'equazione (2). A questo scopo dobbiamo prendere

$$a = -\frac{q}{2}, b = \frac{1 - p}{2}, R^2 = \frac{q^2}{4} + \frac{(1 - p)^2}{4} - r. \qquad (3)$$

Nell'ultima formula (3), in generale, R^2 può risultare negativo. Tuttavia, nel caso in cui l'equazione (2) abbia anche una sola radice reale x_1, vale la seguente uguaglianza

$$x_1^4 + (1 - 2b)\, x_1^2 - 2ax_1 + a^2 + b^2 - R^2 = 0. \qquad (4)$$

Denotando x_1^2 con y_1, l'equazione (4) può essere riscritta come

$$x_1^2 + y_1^2 - 2ax_1 - 2by_1 + a^2 + b^2 - R^2 = 0$$

o come

$$(x_1 - a)^2 + (y_1 - b)^2 = R^2.$$

Quindi, nel caso in cui l'equazione (2) abbia una radice reale, il numero $R^2 = [(1 - p)^2 + q^2]/4-r$ è non negativo, l'equazione

$$(x - a)^2 + (y - b)^2 = R^2$$

è l'equazione di un cerchio, e tutte le radici reali dell'equazione (2) sono le ascisse dei punti di intersezione della parabola $y = x^2$ con questo cerchio. (Nel caso $r = 0$, $R^2 = a^2 + b^2$, questo cerchio passa per l'origine.)

Pertanto, se vengono forniti i coefficienti p, q, r dell'equazione (2) ed è necessario trovare a, b e R^2 con le formule (3), allora se $R^2 < 0$, è noto che l'equazione (2) non ha radici reali. Ma, se $R^2 \geqslant 0$ allora le ascisse dei punti di intersezione della circonferenza con centro (a, b) e raggio R con la parabola $y = x^2$ (disegnato una volta per tutte) fornisce tutte le radici reali dell'equazione (2); ed anche nel caso $R^2 < 0$, il cerchio risultante non può intersecare la parabola e l'equazione (2) non ha radici reali.

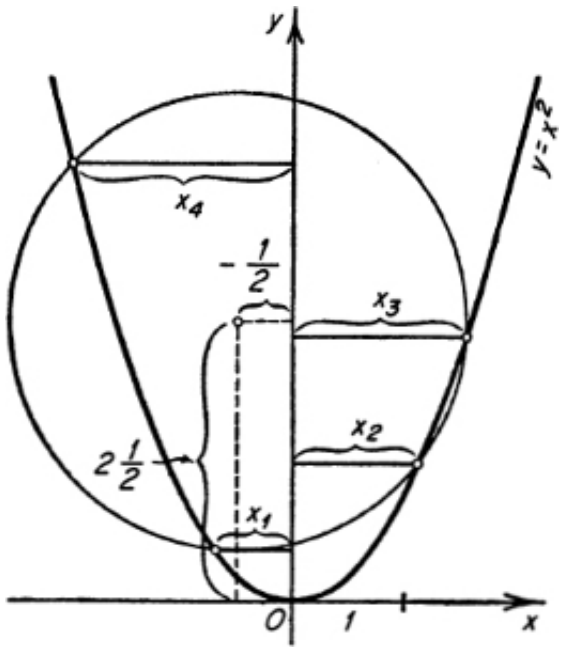


Figura 10.

Esempio. Sia la data equazione di quarto grado:

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

Allora abbiamo

$$a = -\frac{1}{2}, b = \frac{5}{2} = 2\frac{1}{2},$$

$$R = \sqrt{\frac{1}{4} + \frac{25}{4} - \frac{5}{2}} = 2.$$

La Figura 10 mostra il cerchio corrispondente e le radici x_1, x_2, x_3, x_4 dell'equazione data.

I primi quattro argomenti sopra contengono, in una forma abbreviata e un po' più moderna, il contenuto essenziale del libro di Cartesio.

Dai tempi di Cartesio fino ai giorni nostri, la geometria analitica ha subito un immenso sviluppo che è stato molto fruttuoso per molte diverse parti della matematica. Cercheremo nelle sezioni successive di questo capitolo di tracciare le tappe più importanti di questo sviluppo.

Innanzitutto occorre dire che gli inventori dell'analisi infinitesimale erano già in possesso del metodo cartesiano. Che si trattasse di tangenti o normali (perpendicolari alle tangenti nel punto di tangenza) alle curve, o di massimi o minimi di funzioni considerate geometricamente, o del raggio di curvatura di una curva in un punto dato, ecc., si considerava prima l'equazione della curva, col metodo di Cartesio, e poi si trovavano le equazioni della normale, della tangente, ecc. Pertanto l'analisi infinitesimale, cioè il calcolo differenziale e integrale, sarebbe stata inconcepibile senza lo sviluppo preliminare della geometria analitica.

§ 6. La teoria generale dei diametri di Newton

Il primo matematico a compiere un ulteriore grande passo avanti nella stessa geometria analitica fu Newton. Nel 1704 esaminò la teoria delle curve del terzo ordine, cioè delle curve rappresentate da equazioni algebriche di terzo grado in due incognite. Allo stesso tempo trovò, tra le altre cose, un elegante teorema generale sui "diametri", che corrispondono alle secanti in una data direzione. Ha dimostrato quanto segue.

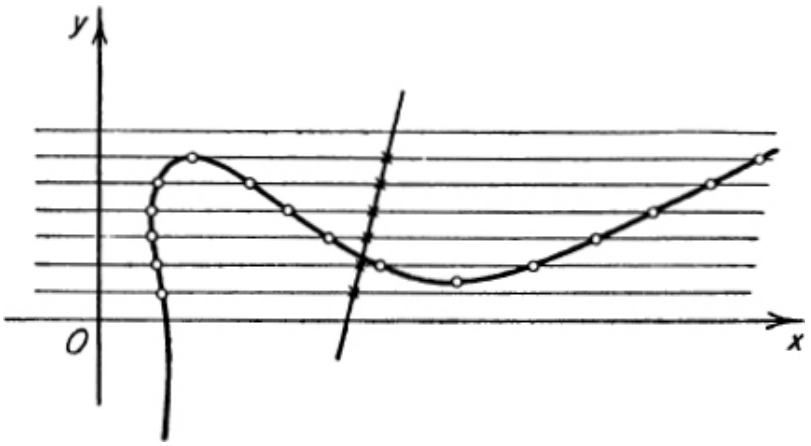


Figura 11.

Sia data una curva di ordine *n*, cioè una curva che sia rappresentata da un'equazione algebrica di grado *n* in due incognite; allora una retta arbitraria che la interseca ha con essa in generale *n* punti comuni. Sia *M* il punto della secante che è il "centro di gravità" di questi punti della

sua intersezione con la curva di ordine n data, cioè il centro di gravità di un insieme di n masse puntiformi situate in questi punti. Risulta che se prendiamo tutti i possibili insiemi di secanti reciprocamente parallele e per ciascuno di essi consideriamo questi centri di massa M, allora per ogni dato insieme di secanti parallele tutti i punti M giacciono su una retta. Newton chiamò questa linea il "diametro" della curva di ordine n corrispondente alla direzione data delle secanti. Poiché la dimostrazione di questo teorema è abbastanza semplice con l'aiuto della geometria analitica, la forniremo qui.

Sia data una curva di ordine n e un insieme di secanti della curva reciprocamente parallele. Scegliamo gli assi delle coordinate in modo che queste secanti siano parallele all'asse Ox (figura 11). Quindi le loro equazioni avranno la forma $y = l$, dove la costante l sarà diversa per le diverse secanti. Sia $F(x, y) = 0$ l'equazione che rappresenta la curva di ordine n rispetto a questi assi di coordinate. È facile mostrare che sotto una trasformazione da un sistema di coordinate rettangolari ad un altro, anche se cambia l'equazione della curva, il suo ordine non cambia (questo sarà mostrato nel §8). Pertanto anche $F(x, y)$ sarà un polinomio di grado n. Per determinare le ascisse dei punti di intersezione della curva con la secante $y = l$, è necessario risolvere le equazioni simultanee $F(x, y) = 0$ e $y = l$; di conseguenza, in generale, si ottiene un'equazione di grado n in x

$$F(x, l) = 0, \qquad (5)$$

da cui troviamo le ascisse $x_1, x_2, \ldots, x_n$. L'ascissa x_c del baricentro dei n punti di intersezione è uguale, per definizione stessa di baricentro, a

$$x_c = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Ma, come è noto dalla teoria delle equazioni algebriche, la somma $x_1 + x_2 + \cdots + x_n$ delle radici di un'equazione è uguale al coefficiente della $(n - 1)$esima potenza dell'incognita x, preso con il segno opposto, diviso per il coefficiente della potenza $n$esima di x. Ma poiché la somma degli esponenti di x e y in ogni termine di $F(x, y)$ è uguale o inferiore a n, il termine in x^n non contiene affatto y ma ha la forma Ax^n, dove A è una costante; e se i termini in x^{n-1} contengono y, lo fanno al massimo alla prima potenza; cioè hanno la forma $x^{n-1}(By + C)$. Di conseguenza, il coefficiente di x^n è A e quello di x^{n-1} è $Bl + C$, e abbiamo per ogni dato l

$$x_c = -\frac{Bl + C}{nA}.$$

Ma la secante è parallela all'asse Ox così che per tutti i suoi punti $y = l$, e quindi anche l'ordinata y del baricentro dei punti della sua intersezione con la data curva di ordine n è uguale a l; si ottiene così alla fine $nAx_c + By_c + C = 0$, cioè le coordinate x_c, y_c dei baricentri di tutte queste secanti soddisfano un'equazione di primo grado, e di conseguenza giacciono su una retta.

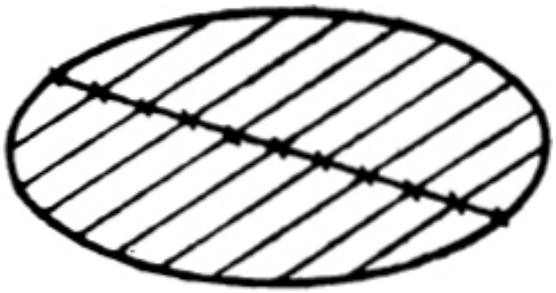


Figura 12.

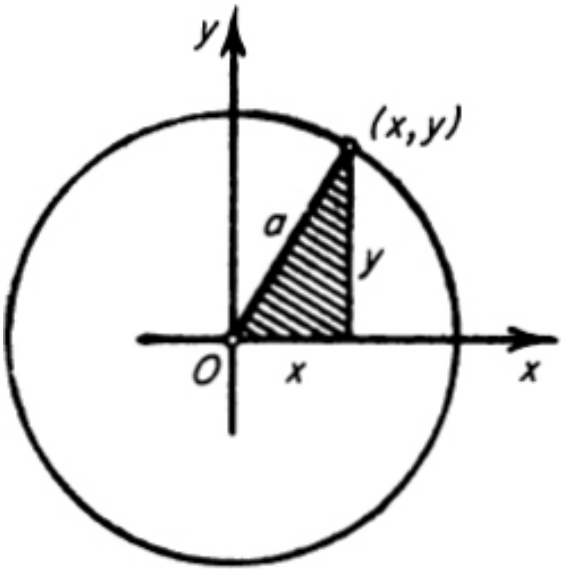


Figura 13.

Il caso in cui $F(x, y)$ non contiene x^n può essere studiato in modo analogo.

Nel caso in cui la curva sia del 2° ordine ($n = 2$) il baricentro di due punti è semplicemente il punto medio tra loro, così che il luogo dei punti medi di corde parallele di una curva del secondo ordine è una retta (figura 12), risultato che per l'ellisse, così come per l'iperbole e la parabola, era già ben noto agli antichi. Ma ciò fu da loro dimostrato, sia pure solo per questi casi parziali, con argomenti geometrici alquanto difficili, e qui un nuovo teorema generale, sconosciuto agli antichi, viene dimostrato in modo del tutto semplice.

Tali esempi rivelano il potere della geometria analitica.

§ 7. Ellisse, iperbole e parabola

In questo paragrafo e nei successivi considereremo le curve del secondo ordine. Prima di studiare l'equazione generale di secondo grado, è utile esaminare alcune delle sue forme più semplici.

Equazione della circonferenza con centro nell'origine. Innanzitutto consideriamo l'equazione

$$x^2 + y^2 = a^2.$$

Rappresenta evidentemente un cerchio con centro nell'origine e raggio a, come segue dal teorema di Pitagora applicato al triangolo rettangolo ombreggiato (figura 13), poiché qualunque punto (x, y) di questo cerchio venga preso, le sue coordinate x e y soddisfano questa equazione, e viceversa, se le coordinate x, y di un punto soddisfano l'equazione, allora il punto appartiene alla circonferenza; cioè il cerchio è l'insieme di tutti quei punti del piano che soddisfano l'equazione.

L'equazione di un'ellisse e la sua proprietà focale. Siano dati due punti F_1 e F_2, la cui distanza è uguale a $2c$. Troveremo l'equazione del luogo di tutti i punti M del piano; la cui somma delle distanze dai punti F_1 e F_2 è pari ad una costante $2a$ (dove, ovviamente, a è maggiore di c). Tale curva è chiamata ellisse e i punti F_1 e F_2 sono i suoi fuochi.

Scegliamo un sistema di coordinate rettangolari tale che i punti F_1 e F_2 giacciono sull'asse Ox e l'origine sia a metà strada tra loro. Quindi le coordinate dei punti F_1 e F_2 saranno $(c, 0)$ e $(-c, 0)$. Prendiamo un punto arbitrario M di coordinate (x, y), appartenente al luogo in questione, e scriviamo che la somma delle sue distanze dai punti F_1 e F_2 è uguale a $2a$,

$$\sqrt{(x-c)^2 + (y-0)^2} + \sqrt{(x+c)^2 + (y-0)^2} = 2a. \qquad (6)$$

Questa equazione è soddisfatta dalle coordinate (x, y) di qualsiasi punto del luogo in esame. Ovviamente è vero anche il viceversa, cioè che qualunque punto le cui coordinate soddisfano l'equazione (6) appartiene a questo luogo. L'equazione (6) è quindi l'equazione del luogo. Resta da semplificarla.

Elevando entrambi i membri alla seconda potenza si ottiene

$$x^2 - 2cx + c^2 + y^2 + 2\sqrt{(x^2 - 2cx + c^2 + y^2)(x^2 + 2cx + c^2 + y^2)} + x^2 + 2cx + c^2 + y^2 = 4a^2,$$

o dopo la semplificazione

$$x^2 + y^2 + c^2 - 2a^2 = -\sqrt{(x^2 + y^2 + c^2)^2 - 4c^2x^2}.$$

Quadrando nuovamente entrambi i membri, otteniamo

$$(x^2 + y^2 + c^2)^2 - 4a^2(x^2 + y^2 + c^2) + 4a^4 = (x^2 + y^2 + c^2)^2 - 4c^2x^2$$

o dopo la semplificazione

$$(a^2 - c^2)x^2 + a^2y^2 = (a^2 - c^2)a^2.$$

Poniamo $a^2 - c^2 = b^2$ (come si può fare poiché $a > c$); quindi otteniamo $b^2x^2 + a^2y^2 = a^2b^2$, e dividendo per a^2b^2 abbiamo

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \qquad (7)$$

Le coordinate (x, y) di qualsiasi punto M del luogo soddisfano quindi l'equazione (7).

Si può invece dimostrare che se le coordinate di un punto soddisfano l'equazione (7) allora soddisfano anche l'equazione (6). Di conseguenza, l'equazione (7) è l'equazione di questo luogo, cioè l'equazione dell'ellisse (figura 14).

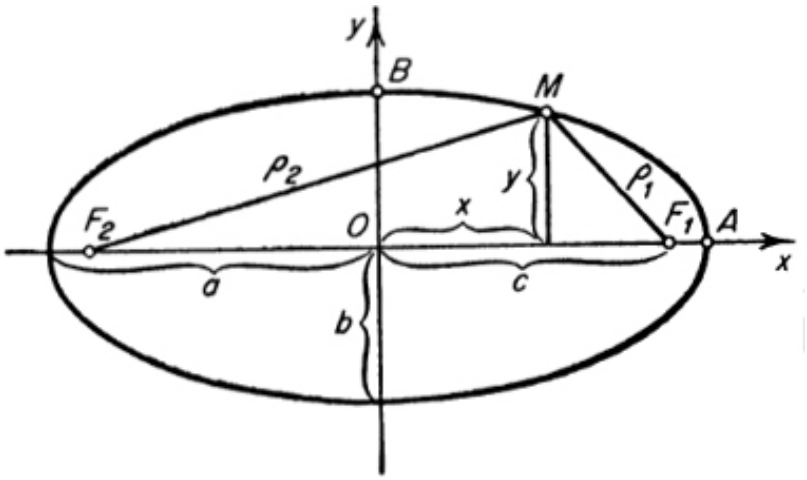


Figura 14.

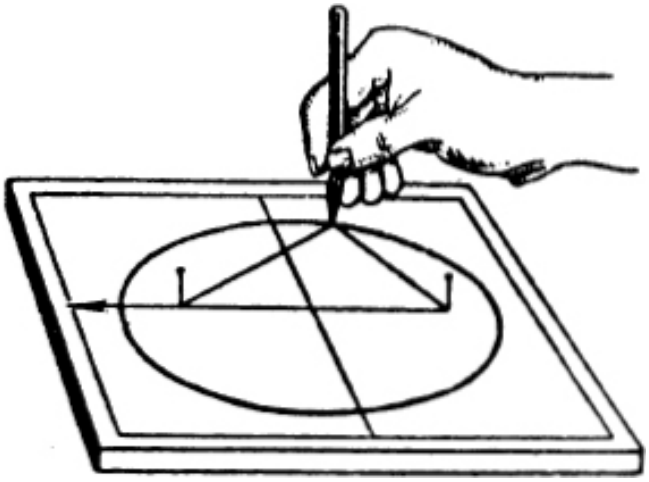


Figura 15.

Questo argomento è un classico esempio di come trovare l'equazione di una curva data da alcune delle sue proprietà geometriche.

Il noto metodo per tracciare un'ellisse mediante un filo (figura 15) si basa sulla proprietà dell'ellisse secondo cui la somma delle distanze di uno qualsiasi dei suoi punti da due punti dati è costante.

Osservazione. Per determinare un'ellisse avremmo potuto prendere, al posto della proprietà focale qui considerata, qualsiasi altra proprietà geometrica che la caratterizza, ad esempio che l'ellisse sia il risultato di una "contrazione uniforme" di un cerchio verso uno dei suoi diametri o qualsiasi altra proprietà.

Sostituendo $y = 0$ nell'equazione (7) dell'ellisse, otteniamo $x = \pm a$, cioè a è la lunghezza del segmento OA (vedi figura 14), chiamato semiasse maggiore dell'ellisse. Analogamente, sostituendo $x = 0$, si ottiene $y = \pm b$, cioè b è la lunghezza del segmento OB, detto semiasse minore dell'ellisse.

Il numero $\varepsilon = c/a$ è chiamato eccentricità dell'ellisse, per cui, essendo $c = \sqrt{a^2 - b^2} < a$, l'eccentricità di un'ellisse è minore di 1. Nel caso di un cerchio, $c = 0$ e di conseguenza $\varepsilon = 0$; entrambi i fuochi sono in un punto, il centro del cerchio (poiché $OF_1 = OF_2 = 0$), ma il metodo precedente di disegnare la curva con un filo è ancora valido.

Leggi del moto planetario. Studiando le lunghe osservazioni di Tycho Brahe sul movimento del pianeta Marte, Keplero scoprì che i pianeti ruotano attorno al Sole lungo ellissi tali che il Sole occupa un fuoco dell'ellisse (l'altro fuoco rimane non occupato e non gioca alcun ruolo nel movimento di un pianeta attorno al Sole) (figura 16). Keplero osservò anche che il raggio focale ρ in tempi uguali spazza settori di uguale area,[2] e Newton dimostrò che la necessità di un tale movimento consegue matematicamente dalla legge del moto (proporzionalità dell'accelerazione alla forza) e dalla legge di gravitazione universale.

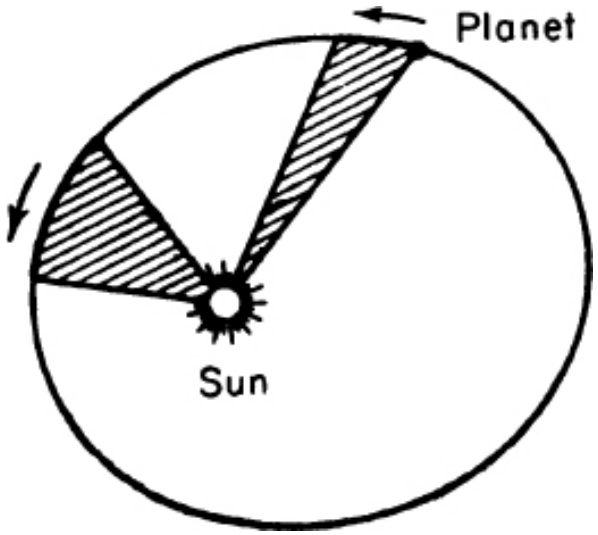


Figura 16. «Planet»: pianeta; «Sun»: Sole.

L'ellisse dell'inerzia. Come esempio di applicazione dell'ellisse in un problema tecnico, consideriamo la cosiddetta ellisse d'inerzia di una piastra.

Supponiamo che la piastra abbia uno spessore uniforme e sia costituita da materiale omogeneo, ad esempio una piastra di zinco di forma arbitraria. La ruotiamo attorno ad un asse nel suo piano. Un corpo in moto rettilineo possiede, come è noto, un'inerzia rispetto a tale moto rettilineo proporzionale alla sua massa (indipendentemente dalla forma del corpo e dalla distribuzione della massa). Allo stesso modo, un corpo che ruota attorno ad un asse, ad esempio un volano, presenta inerzia rispetto a questa rotazione. Ma nel caso della rotazione, l'inerzia non è solo proporzionale alla massa del corpo rotante ma dipende anche dalla distribuzione della massa del corpo rispetto all'asse di rotazione, poiché l'inerzia rispetto alla rotazione è maggiore se la massa è più lontana dall'asse. Ad esempio, è molto semplice portare subito un bastone in rapida rotazione attorno al suo asse longitudinale (figura 17a). Ma se proviamo a portarlo subito in rotazione veloce attorno ad un asse perpendicolare alla sua lunghezza, anche se l'asse passa per il suo punto medio, troveremo che, a meno che questo bastone non sia molto leggero, dobbiamo esercitare uno sforzo considerevole (figura 17b).



Figura 17a.

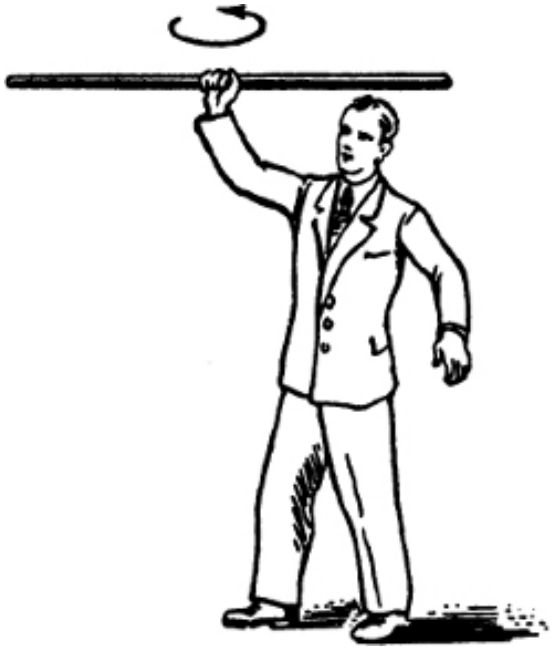


Figura 17b.

È possibile dimostrare che l'inerzia di un corpo rispetto alla rotazione attorno ad un asse, il cosiddetto momento d'inerzia del corpo rispetto all'asse, è pari a $\Sigma r_i^2 m_i$ (dove per $\Sigma r_i^2 m_i$ intendiamo la somma $r_1^2 m_1 + r_2^2 m_2 + \cdots + r_n^2 m_n$ e pensiamo al corpo scomposto in piccolissimi elementi, con m_i la massa del *i*esimo elemento e r_i la distanza dell'elemento *i*esimo dall'asse di rotazione, essendo la sommatoria di tutti gli elementi del corpo).

Torniamo alla nostra piastra. Sia *O* (figura 18) un punto di questa piastra. Consideriamo i momenti di inerzia J_u della piastra rispetto ad un asse

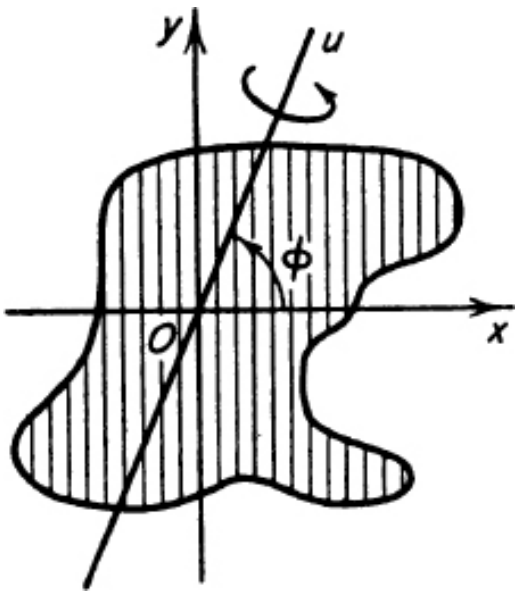


Figura 18.

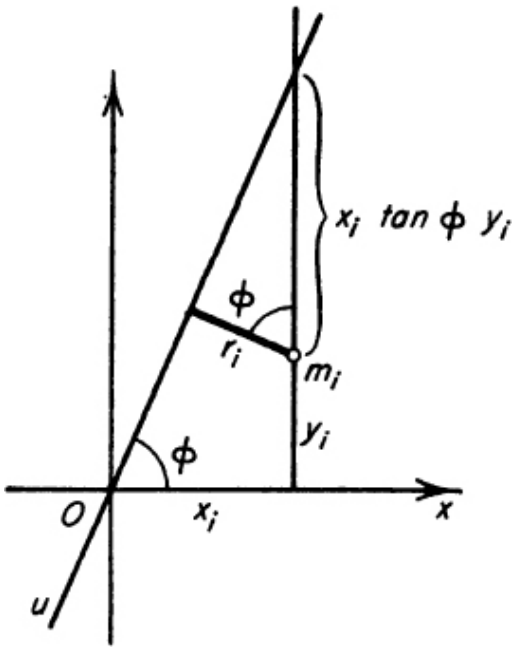


Figura 19.

di rotazione *u* passante per *O* e giacente nel piano della piastra. A questo scopo prendiamo il punto *O* come origine di un sistema di coordinate cartesiane e scegliamo gli assi arbitrari Ox e Oy nel piano della piastra; poi caratterizzeremo l'asse di rotazione *u* con l'angolo ϕ che forma con l'asse Ox. È facile vedere (figura 19) che

$$r_i = |(x_i \tan\phi - y_i)\cos\phi| = |x_i \sin\phi - y_i \cos\phi|.$$

Quindi

$$\Sigma r_i^2 m_i = \Sigma (x_i^2 \sin^2\phi - 2x_i y_i \sin\phi\cos\phi + y_i^2 \cos^2\phi)\, m_i$$
$$= \sin^2\phi\ \Sigma x_i^2 m_i - 2\sin\phi\cos\phi\ \Sigma x_i y_i m_i + \cos^2\phi\ \Sigma y_i^2 m_i\ .$$

Le quantità $\sin^2\phi$, 2 sin ϕ cos ϕ e $\cos^2\phi$ vengono prese al di fuori del segno di sommatoria, poiché sono costanti per un dato asse *u*. Ora scriviamo

$$\Sigma x_i^2 m_i = A, \quad -\Sigma x_i y_i m_i = B, \quad \Sigma y_i^2 m_i = C.$$

Le quantità *A*, *B* e *C* non dipendono dalla scelta dell'asse *u*, ma solo dalla forma della piastra, dalla distribuzione della sua massa e dalla scelta fissa degli assi coordinati Ox e Oy. Di conseguenza,

$$J_u = A\sin^2\phi + 2B\sin\phi\cos\phi + C\cos^2\phi.$$

Consideriamo tutti i possibili assi *u* nel piano della piastra passanti per il punto *O* e riportiamo su ciascuno di essi, a partire da *O*, una lunghezza ρ uguale all'inverso della radice quadrata del momento d'inerzia J_u della piastra rispetto a quell'asse, cioè $\rho = 1/\sqrt{J_u}$. Otteniamo allora

$$\frac{1}{\rho^2} = A \sin^2 \phi + 2B \sin \phi \cos \phi + C \cos^2 \phi.$$

Ma

$$x = \rho \cos \phi, \quad y = \rho \sin \phi,$$

per cui l'equazione di questo luogo ha la seguente forma:

$$Cx^2 + 2Bxy + Ay^2 = 1.$$

Si ottiene una curva del secondo ordine evidentemente limitata e chiusa; cioè è un'ellisse (figura 20), poiché tutte le altre curve del secondo ordine, come mostreremo più avanti, o sono infinite o si riducono a un punto.

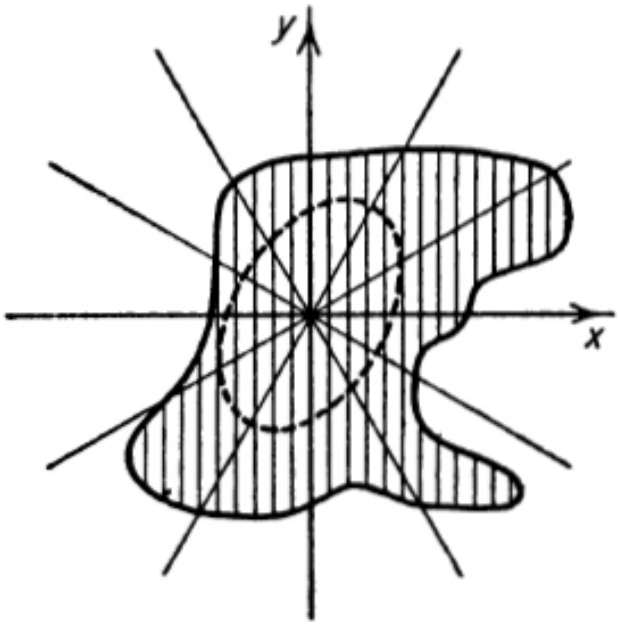


Figura 20.

Si ottiene così un risultato notevole: qualunque siano la forma e le dimensioni della piastra e la distribuzione della sua massa, il suo momento d'inerzia rispetto ai diversi assi giacenti nel suo piano e passanti per un punto dato *O* è caratterizzato da una certa ellisse. Più precisamente, è caratterizzata da questa ellisse la quantità ρ, inversamente proporzionale alla radice quadrata del momento d'inerzia. Tale curva si chiama ellisse d'inerzia della piastra rispetto al punto *O*. Se *O* è il baricentro della piastra, si parla di ellisse centrale d'inerzia.

L'ellisse d'inerzia gioca un ruolo importante nella meccanica; in particolare ha un'importante applicazione nella resistenza dei materiali. Nella teoria della resistenza dei materiali si dimostra che la resistenza alla flessione di una trave di data sezione trasversale è proporzionale al momento d'inerzia della sua sezione trasversale rispetto all'asse passante per il baricentro della sezione trasversale e perpendicolare alla direzione della forza di flessione. Chiariamo questo con un esempio. Supponiamo che un ponte su un ruscello sia costituito da una tavola che si piega sotto il peso di un pedone che vi passa sopra. Se la stessa tavola (non

più spessa di prima) viene posta “sul bordo”, non si piega quasi per niente, cioè una tavola posizionata sul bordo è, per così dire, più resistente. Ciò deriva dal fatto che il momento d'inerzia della sezione trasversale della tavola (ha la forma di un rettangolo allungato che possiamo pensare uniformemente ricoperto di massa) è maggiore rispetto all'asse perpendicolare al suo lato lungo che rispetto all'asse parallelo al suo lato lungo. Se poniamo la tavola non esattamente piana né di spigolo ma obliquamente, o anche se non prendiamo affatto una tavola ma un'asta di sezione trasversale arbitraria, ad esempio una rotaia, la resistenza alla flessione sarà comunque proporzionale al momento d'inerzia della sua sezione trasversale rispetto all'asse corrispondente. La resistenza di una trave alla flessione è quindi caratterizzata dall'ellisse d'inerzia della sua sezione trasversale.

Per una trave rettangolare ordinaria questa ellisse avrà la forma mostrata in figura 21. La rigidità di tale trave sotto un carico nella direzione dell'asse Oz è proporzionale a bh^3.

Le travi in acciaio hanno spesso una sezione trasversale a forma di ⌡; per tali travi la sezione trasversale e l'ellisse d'inerzia sono rappresentate in figura 22.

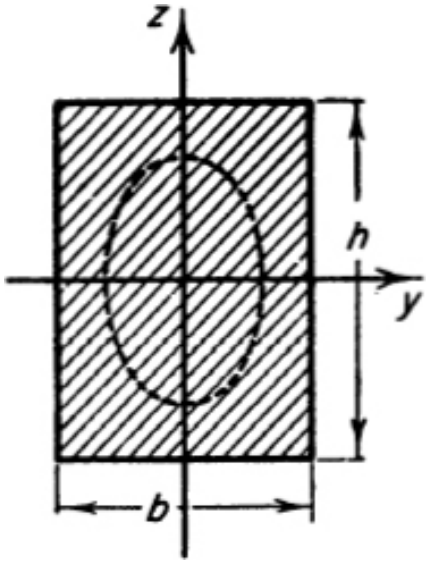


Figura 21.

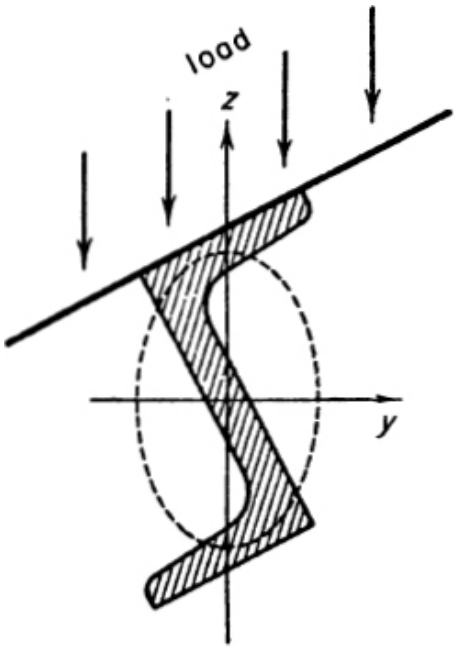


Figura 22. «load»: carico.

La massima resistenza alla flessione è nella direzione *z*. Quando queste travi vengono utilizzate, ad esempio come travi del tetto sotto il carico di neve e il proprio peso, agiscono direttamente contro la flessione in una direzione vicina a questa direzione più vantaggiosa.

L'iperbole e le sue proprietà focali. Consideriamo ora l'equazione

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

rappresenta una curva chiamata iperbole. Se indichiamo con *c* un numero tale che $c^2 = a^2 + b^2$, allora è possibile dimostrare che un'iperbole è il luogo di tutti i punti la differenza delle cui distanze dai punti F_1 e F_2 sull'asse Ox con ascisse *c* e *-c* è una costante: $\rho_2 - \rho_1 = 2a$ (figura 23). I punti F_1 e F_2 sono detti fuochi.

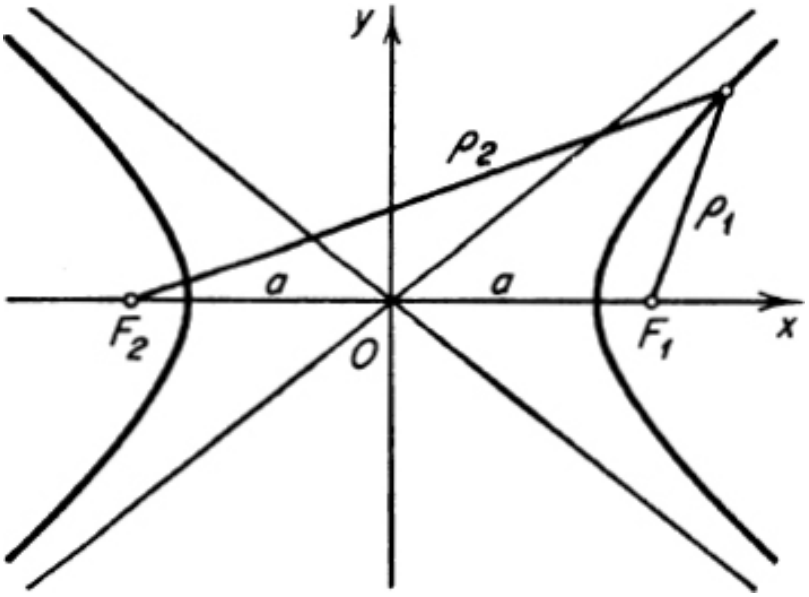


Figura 23.

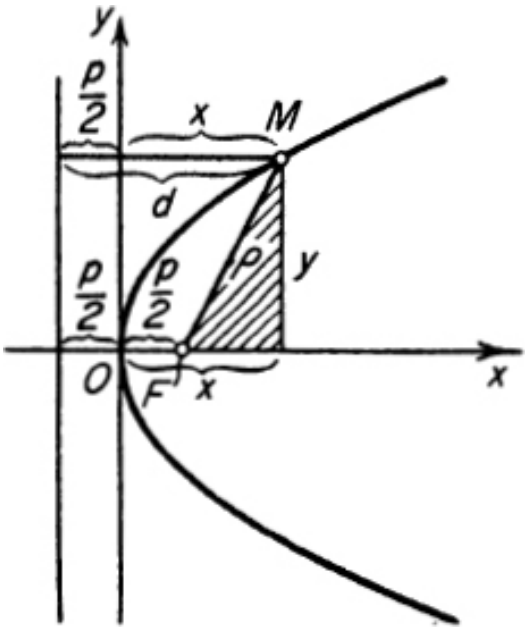


Figura 24.

La parabola e la sua direttrice. Infine, consideriamo l'equazione

$$y^2 = 2px$$

e chiamiamo parabola la curva corrispondente. Il punto F giacente sull'asse Ox con ascissa $p/2$ è detto fuoco della parabola, e la retta $x = -p/2$, parallela all'asse Oy, è la sua direttrice. Sia M un punto qualsiasi della parabola (figura 24), ρ la lunghezza del suo raggio focale MF, e d la lunghezza della perpendicolare caduta da essa alla direttrice. Calcoliamo ρ e d per il punto M. Dal triangolo ombreggiato otteniamo $\rho^2 = (x - p/2)^2 + y^2$. Poiché il punto M giace sulla parabola, abbiamo $y^2 = 2px$, quindi[E1]

$$\rho^2 = \left(x - \frac{p}{2}\right)^2 + 2px = \left(x + \frac{p}{2}\right)^2.$$

Ma direttamente dalla figura risulta chiaro che $d = x + p/2$. Pertanto $\rho^2 = d^2$, cioè $\rho = d$. Il ragionamento inverso mostra che se per un dato punto abbiamo $\rho = d$, allora il punto giace sulla parabola. Pertanto una parabola è il luogo dei punti equidistanti da un dato punto F (chiamato fuoco) e una data retta d (chiamata direttrice).

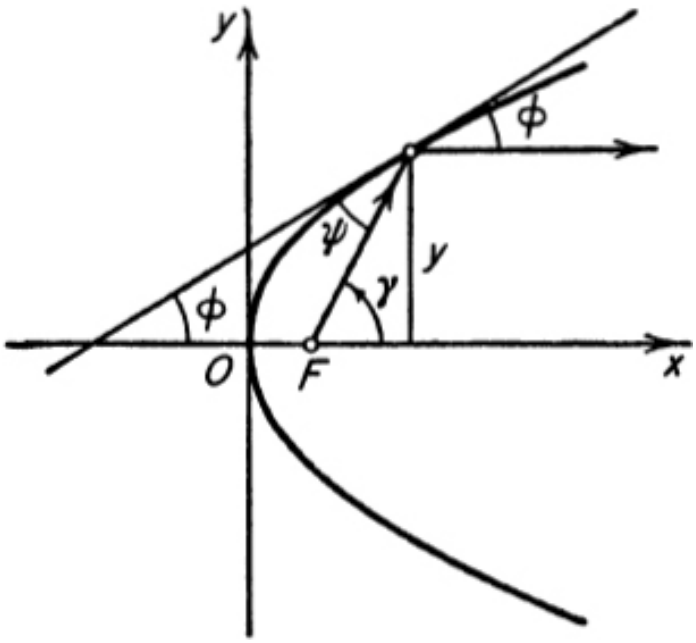


Figura 25.

La proprietà della tangente ad una parabola. Esaminiamo un'importante proprietà della tangente ad una parabola e la sua applicazione in ottica. Poiché per una parabola $y^2 = 2px$ abbiamo $2y\,dy = 2p\,dx$, ne consegue che la derivata, ovvero la pendenza della tangente, è uguale a $dy/dx = \tan\phi = p/y$ (figura 25).

D'altra parte, dalla figura segue direttamente che

$$\tan\gamma = \frac{y}{x - p/2}.$$

Ma

$$\tan 2\phi = \frac{2\,p/y}{1 - p^2/y^2} = \frac{2py}{y^2 - p^2} = \frac{2py}{2px - p^2} = \frac{y}{x - p/2},$$

cioè $\gamma = 2\phi$, e poiché $\gamma = \phi + \psi$, quindi $\psi = \phi$. Di conseguenza, in virtù della legge (l'angolo di incidenza è uguale all'angolo di riflessione) un fascio di luce, partito dal fuoco F e riflesso da un elemento della parabola (la cui direzione coincide con la direzione della tangente) viene riflesso parallelamente all'asse Ox, cioè parallelo all'asse di simmetria della parabola.

Su questa proprietà della parabola si basa la costruzione dei telescopi riflettori, inventati da Newton. Se realizziamo uno specchio concavo la cui superficie è un cosiddetto paraboloide di rivoluzione, cioè una superficie ottenuta dalla rotazione di una parabola attorno al suo asse di simmetria, allora tutti i raggi luminosi provenienti da un punto qualsiasi di un corpo celeste giacente rigorosamente nella direzione dell'"asse" dello specchio vengono raccolti dallo specchio (figura 26) in un punto, cioè nel suo fuoco. I raggi provenienti da qualche altro punto del corpo celeste, non essendo esattamente paralleli all'asse dello specchio, si raccolgono quasi in un punto nelle vicinanze del fuoco. Così, nel cosiddetto piano focale passante per il fuoco dello specchio e perpendicolare al suo asse, si ottiene l'immagine inversa della stella; quanto più questa immagine è lontana dal fuoco, tanto più diffusa sarà, poiché solo i raggi esattamente paralleli all'asse dello specchio vengono raccolti dallo specchio in un punto. L'immagine così ottenuta può essere visionata in un apposito microscopio, il cosiddetto

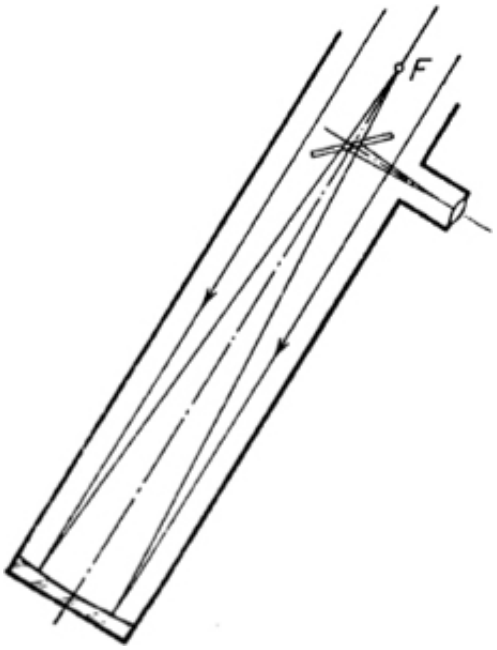


Figura 26.

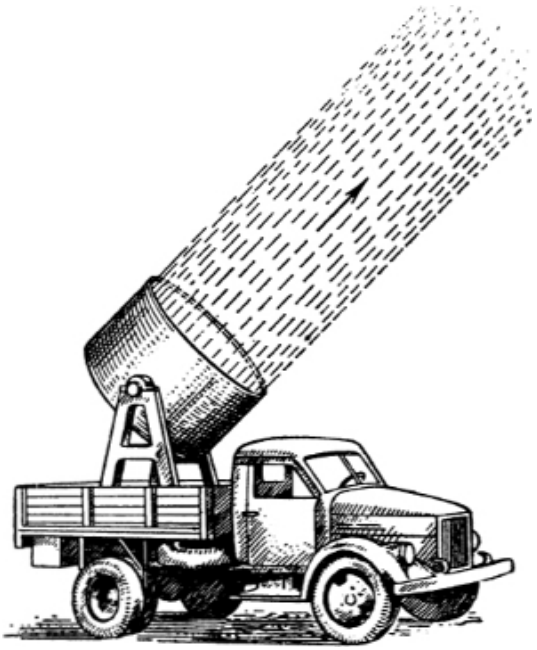


Figura 27.

oculare del telescopio, direttamente oppure, per non tagliare la luce della stella con la propria testa, dopo la riflessione in un piccolo specchio piano, fissato al telescopio vicino al fuoco (un po' più vicino del fuoco allo specchio concavo) con un angolo di 45°.

Il proiettore (figura 27) si basa sulla stessa proprietà della parabola. In esso, al contrario, una forte sorgente luminosa è posta al fuoco di uno specchio paraboloidale, in modo che i suoi raggi siano riflessi dallo specchio in un fascio parallelo al suo asse. I fari delle automobili sono costruiti in modo simile (figura 28).

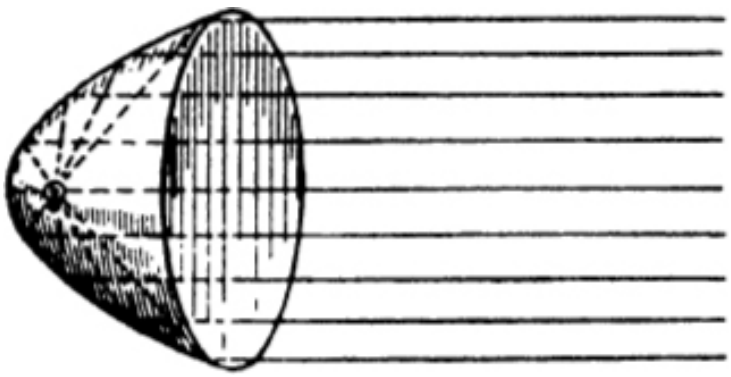


Figura 28.

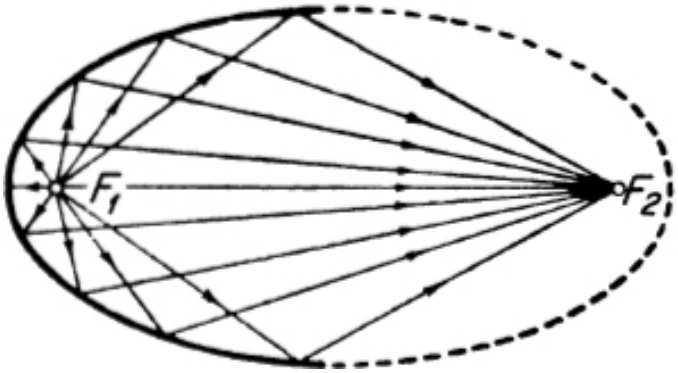


Figura 29.

Nel caso di un'ellisse, come è facile mostrare, i raggi uscenti da uno dei suoi fuochi F_1 e riflessi dall'ellisse si raccolgono nell'altro fuoco F_2 (figura 29), e nell'iperbole i raggi originati da uno dei suoi fuochi F_1 vengono riflessi da esso come se provenissero dall'altro fuoco F_2 (figura 30).

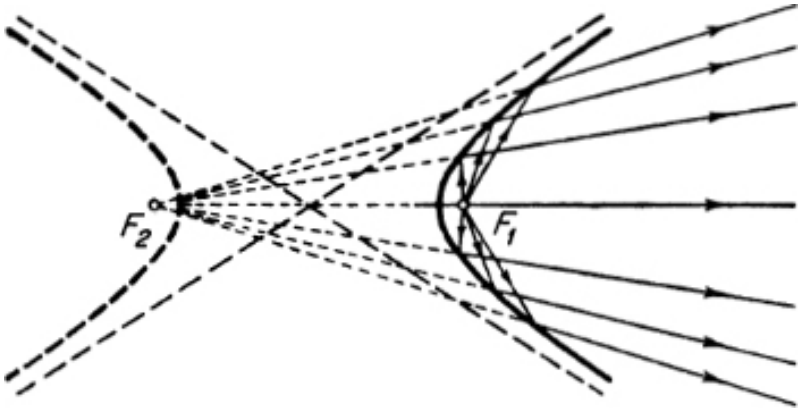


Figura 30.

Le direttrici dell'ellisse e dell'iperbole. Come la parabola, l'ellisse e l'iperbole hanno direttrici, in questo caso due ciascuna. Se consideriamo un fuoco e la direttrice “dalla sua stessa parte”, allora per tutti i punti M dell'ellisse abbiamo $\rho/d =$ € dove la costante € è l'eccentricità, che per un'ellisse è sempre minore di 1; e per tutti i punti del ramo corrispondente dell'iperbole abbiamo anche $\rho/d =$ €, dove € è ancora l'eccentricità, che per un'iperbole è sempre maggiore di 1.

Pertanto l'ellisse, la parabola e un ramo dell'iperbole sono i luoghi di tutti quei punti del piano per i quali è costante il rapporto tra la loro distanza ρ dal fuoco e la loro distanza d dalla direttrice (figure 31 e 32). Per l'ellisse questa costante è minore dell'unità, per la parabola è uguale all'unità e per l'iperbole è maggiore dell'unità.

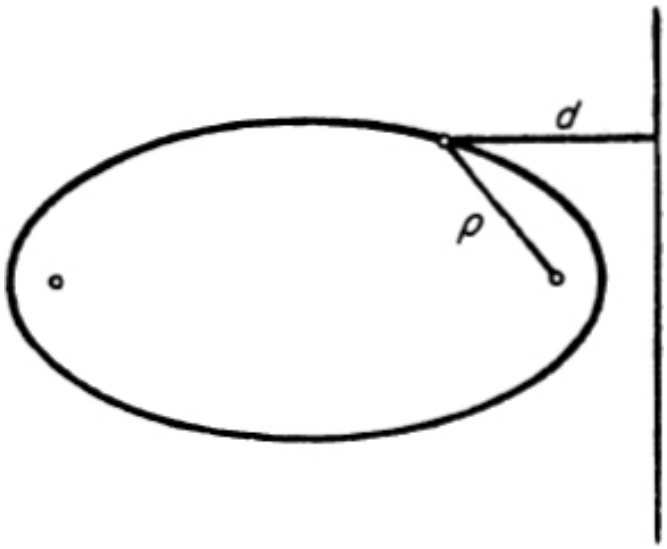


Figura 31.

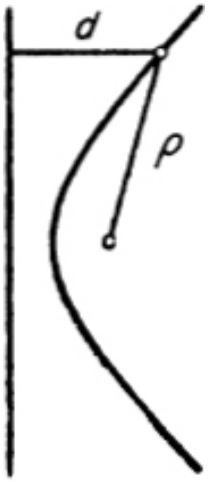


Figura 32.

In questo senso la parabola è il caso “limitante” o di “transizione” dall'ellisse all'iperbole.

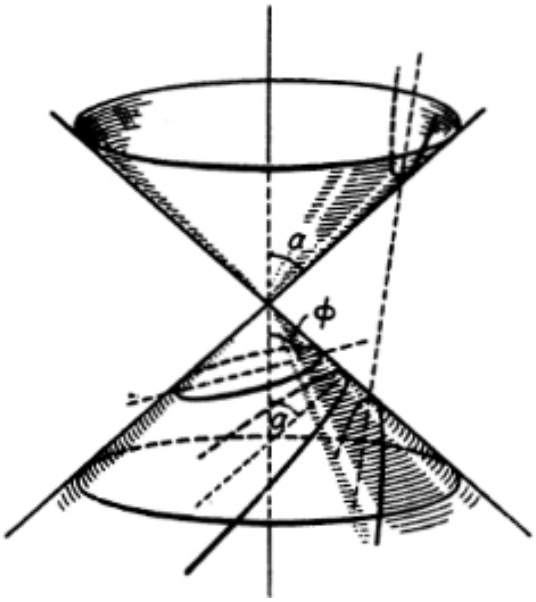


Figura 33.

Sezioni coniche. Già gli antichi greci avevano studiato in dettaglio le curve ottenute intersecando un cono circolare retto con un piano. Se il piano intersecante forma con l'asse del cono un angolo ϕ di 90°, cioè è ad esso perpendicolare, allora la sezione ottenuta è un cerchio. È facile dimostrare che se l'angolo ϕ è minore di 90°, ma maggiore dell'angolo α che formano le generatrici del cono con il proprio asse, allora si ottiene un'ellisse. Se ϕ è uguale a α, risulta una parabola e se ϕ è minore di α allora otteniamo un'iperbole come sezione (figura 33).

La parabola come grafico della proporzionalità quadratica e l'iperbole come grafico della proporzionalità inversa. Ricordiamo che il grafico della proporzionalità quadratica

$$y = kx^2$$

è una parabola (figura 34) e cioè il grafico di proporzionalità inversa

$$y = \frac{k}{x} \quad \text{or} \quad xy = k$$

è un'iperbole (figura 35), come dimostreremo facilmente in seguito. Un'iperbole è stata definita in precedenza come la curva rappresentata dall'equazione

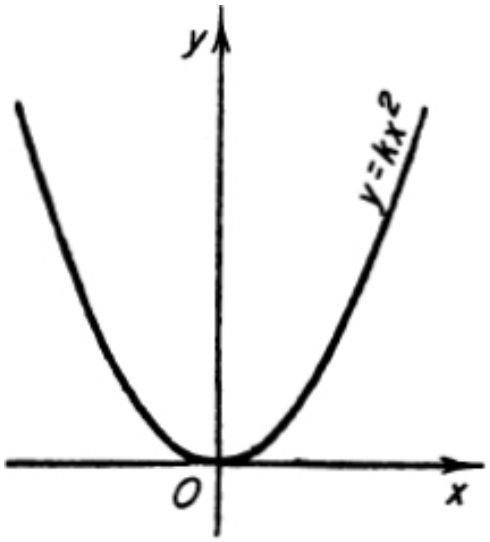


Figura 34.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Nel caso speciale $a = b$ la cosiddetta iperbole equilatera gioca tra le iperboli lo stesso ruolo che il cerchio gioca tra le ellissi. In questo caso, se ruotiamo gli assi delle coordinate di 45° (figura 36) l'equazione nelle nuove coordinate (x', y') avrà la forma

$$x'y' = k.$$

Abbiamo ora considerato tre importanti curve del secondo ordine: l'ellisse, l'iperbole e la parabola, e per le loro definizioni abbiamo preso le cosiddette equazioni canoniche

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{and} \quad y^2 = 2px,$$

con cui sono rappresentate.

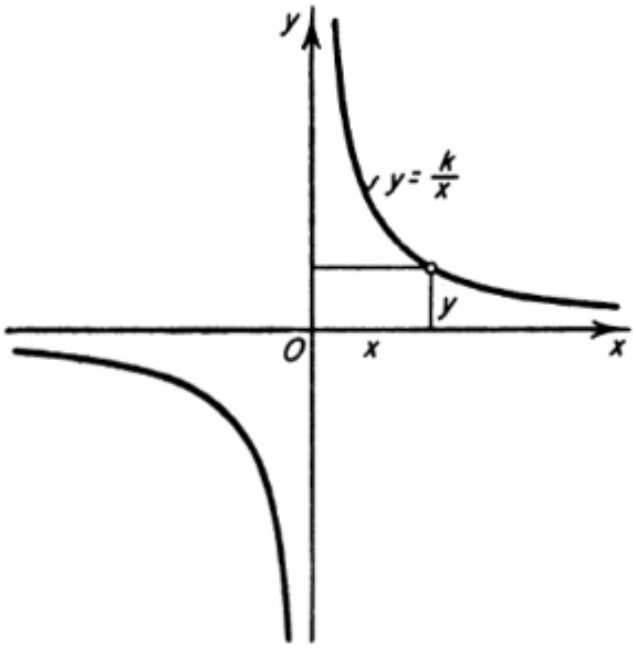


Figura 35.

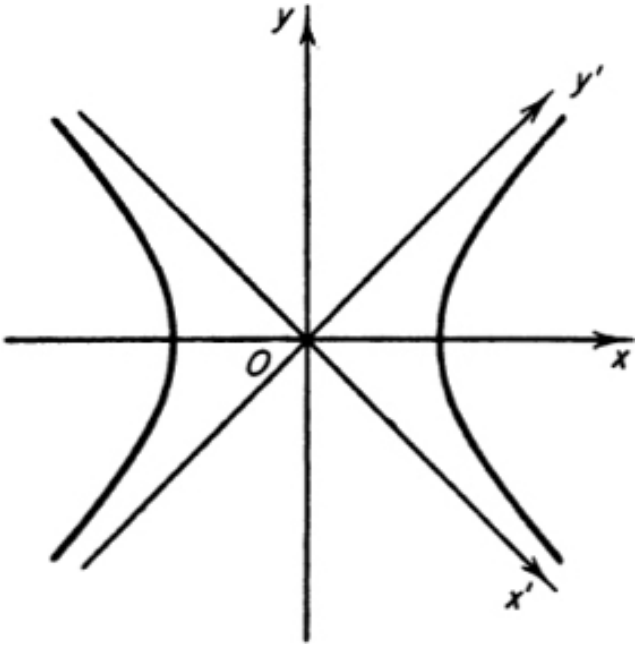


Figura 36.

Passiamo ora allo studio dell'equazione generale di secondo grado in due incognite, cioè alla domanda su quali tipi di curve siano rappresentate da questa equazione.

§ 8. Riduzione dell'equazione generale di secondo grado alla forma canonica

La prima presentazione coerente della geometria analitica da parte di Eulero. Un passo significativo nello sviluppo della geometria analitica fu la pubblicazione nel 1748 del libro "Introduzione all'analisi", nel secondo volume del quale, tra l'altro relativo alla teoria delle funzioni e ad altri rami dell'analisi, per la prima volta fu data una presentazione della geometria analitica nel piano con uno studio dettagliato delle curve del secondo ordine, molto vicino a quello dato nei libri di testo contemporanei di geometria analitica, e anche con uno studio delle curve di ordine superiore. Questa fu la prima esposizione della geometria analitica nel senso contemporaneo del termine.

La nozione di riduzione di un'equazione alla forma canonica. Un'equazione di secondo grado[3]

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

contiene sei termini, non tre o solo due come nelle equazioni canoniche sopra dell'ellisse, dell'iperbole e della parabola. Questo non perché tale equazione rappresenti una curva più complicata, ma perché il sistema di coordinate potrebbe non essere adatto ad essa. Risulta che se scegliamo un sistema di coordinate cartesiane adatto, allora un'equazione di secondo grado con due variabili può sempre essere ridotta a una delle seguenti forme canoniche:

1. $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0.$  Ellisse
2. $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0.$  Ellisse immaginaria
3. $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0.$ Punto (coppia di rette immaginarie che si intersecano in un punto reale)
4. $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0.$  Iperbole
5. $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0.$  Coppia di rette incidenti
6. $y^2 - 2px = 0.$  Parabola

7. $x^2 - a^2 = 0.$  Coppia di rette parallele

8. $x^2 + a^2 = 0.$  Coppia di rette parallele immaginarie

9. $x^2 = 0.$  Coppia di rette coincidenti

dove a, b, p, non sono uguali a zero.

Le equazioni 1, 4 e 6 delle forme canoniche enumerate ci sono già ben note; queste sono le equazioni canoniche dell'ellisse, dell'iperbole e della parabola. Due di esse non sono soddisfatte da alcun punto, vale a dire le equazioni 2 e 8. Infatti, il quadrato di un numero reale è sempre positivo o zero, per cui sul membro sinistro dell'equazione 2 la somma dei termini $x^2/a^2 + y^2/b^2$ non è mai negativa, e poiché compare anche il termine + 1, il membro di sinistra non può essere uguale a zero; analogamente nell'equazione 8, il numero x^2 non è negativo e a^2 è positivo. Da queste considerazioni consegue che solo (x = 0, y = 0) soddisfa l'equazione 3, cioè un punto, l'origine. L'equazione 5 può essere scritta come $(x/a - y/b)(x/a + y/b) = 0$, da cui vediamo che è soddisfatta da quei punti e solo da quei punti per i quali una delle espressioni di primo grado $x/a - y/b$ o $x/a + y/b$ è uguale a zero; quindi la curva che rappresenta è questa coppia di rette che si intersecano. L'equazione 7 dà analogamente $(x - a)(x + a) = 0$; cioè, la curva corrispondente è una coppia di rette parallele $x = a$ e $x = -a$. Infine, la curva 9 è un caso limite speciale della curva 7, quando $a = 0$; cioè, è una coppia di rette coincidenti.

Formule di trasformazioni di coordinate. Per ottenere l'importante risultato indicato sui possibili tipi di curve del secondo ordine, è necessario prima dedurre le formule secondo le quali le coordinate rettangolari dei punti variano al variare del sistema di coordinate.

Siano x, y le coordinate di un punto M relativo agli assi Oxy. Trasliamo questi assi paralleli a se stessi nella posizione $O'x'y'$ e siano le coordinate della nuova origine O' rispetto ai vecchi assi

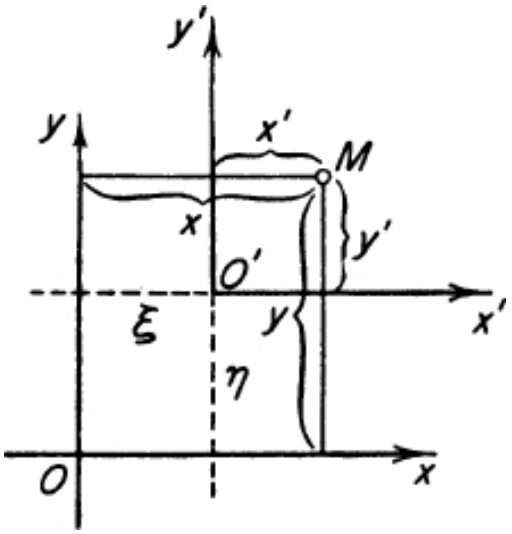


Figura 37.

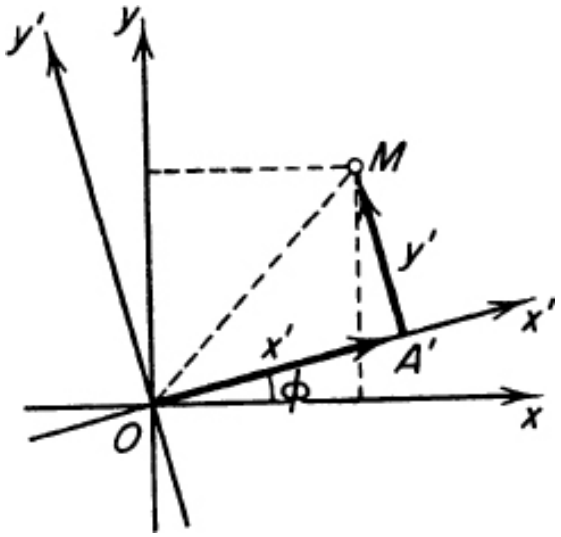


Figura 38.

ξ e η. È evidente (figura 37) che le nuove coordinate x', y' del punto M sono collegate con le sue vecchie coordinate x, y dalle formule

$$x = x' + \xi,$$

$$y = y' + \eta,$$

che sono le formule della cosiddetta traslazione parallela degli assi. Se ruotiamo gli assi originali Oxy attorno all'origine in senso antiorario di un angolo ϕ allora, come è facile vedere (figura 38), se proiettiamo la linea poligonale $OA'M$ composta dai nuovi segmenti di coordinate x', y' rispettivamente sull'asse Ox e sull'asse Oy, otteniamo

$$x = x' \cos\phi - y' \sin\phi,$$

$$y = x' \sin\phi + y' \cos\phi,$$

che sono le formule per la trasformazione delle coordinate sotto la rotazione di un sistema di coordinate rettangolare.

Se ci viene data un'equazione $F(x, y) = 0$ di una curva relativa agli assi Oxy e vogliamo scrivere l'equazione trasformata della stessa curva, cioè relativa ai nuovi assi $O'x'y'$, allora dobbiamo sostituire x e y nell'equazione $F(x, y) = 0$ con le loro espressioni in termini di x' e y', date dalle formule della trasformazione. Ad esempio, con traslazione parallela degli assi, otteniamo l'equazione trasformata

$$F(x' + \xi, y' + \eta) = 0,$$

e sotto rotazione degli assi l'equazione

$$F(x' \cos\phi - y' \sin\phi, x' \sin\phi + y' \cos\phi) = 0.$$

Notiamo che sotto una trasformazione su nuovi assi il grado di un'equazione non cambia. Il grado, infatti, non può aumentare, poiché le formule di trasformazione sono di primo grado. Ma neanche il grado può diminuire, poiché in tal caso la trasformazione inversa delle coordinate lo aumenterebbe (ed è anch'esso di primo grado).

La riduzione di un'equazione generale di secondo grado ad una delle 9 forme canoniche. Mostriamo ora che data una qualsiasi equazione di secondo grado in due incognite possiamo sempre prima ruotare gli assi e poi traslarli parallelamente a se stessi in modo tale che l'equazione trasformata per gli assi finali avrà una delle forme 1, 2, ⋯, 9.

In effetti, sia data un'equazione che di secondo grado abbia la forma

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \qquad (8)$$

Ruotiamo gli assi di un angolo ϕ, che selezioniamo nel modo seguente. Sostituendo x e y nell'equazione (8) con le loro espressioni in termini di nuove coordinate (secondo le formule di rotazione), troviamo, dopo aver raccolto termini simili, che il coefficiente $2B'$ nell'equazione trasformata

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F' = 0$$

è uguale a

$$\begin{aligned} 2B' &= -2A \sin\phi \cos\phi + 2B(\cos^2\phi - \sin^2\phi) + 2C \sin\phi \cos\phi \\ &= 2B \cos 2\phi - (A - C) \sin 2\phi. \end{aligned}$$

Ponendolo uguale a zero, otteniamo $2B \cos 2\phi = (A - C) \sin 2\phi$, da cui

$$\cot 2\phi = \frac{A - C}{2B}.$$

Poiché la cotangente varia da -∞ a +∞, possiamo sempre trovare un angolo ϕ per il quale questa uguaglianza è soddisfatta. Ruotando gli assi di questo angolo, troviamo che per gli assi

ruotati Ox′y′ l'equazione della nostra curva, rappresentata per gli assi iniziali dall'equazione (8), ha la forma

$$A'x'^2 + C'y'^2 + 2D'x' + 2E'y' + F = 0, \qquad (9)$$

cioè che non contenga il termine con il prodotto delle coordinate (F rimane invariato, poiché le formule di rotazione non contengono termini costanti).

Ora trasliamo gli assi già ruotati Ox′ y′ paralleli a se stessi nella posizione $O''\,x''\,y''$, e lasciamo che le coordinate della nuova origine O'' relativa agli assi Ox′y′ siano ξ', η'. L'equazione della nostra curva per questi assi finali sarà

$$A'(x'' + \xi')^2 + C'(y'' + \eta')^2 + 2D'(x'' + \xi') + 2E'(y'' + \eta') + F = 0. \qquad (10)$$

Mostriamo ora che possiamo sempre selezionare ξ' e η' (cioè possiamo traslare gli assi Ox′y′ paralleli a se stessi) in modo tale che l'equazione finale per gli assi $O''\,x''\,y''$ abbia una delle forme canoniche 1, 2, ···,9.

Rimuovendo tutte le parentesi nell'equazione (10) e raccogliendo termini simili, otteniamo

$$A'x''^2 + C'y''^2 + 2(A'\xi' + D')x'' + 2(C'\eta' + E')y'' + F' = 0, \qquad (10')$$

dove abbiamo indicato con F' la somma di tutti i termini costanti; il suo valore al momento non ci interessa.

Consideriamo tre possibili casi.

I. A' e C' entrambi diversi da zero. In questo caso, prendendo $\xi' = -D'/A'$, $\eta' = -E'/C'$, eliminiamo i termini con le prime potenze di x'' e y'' e otteniamo un'equazione della forma

$$A'x''^2 + C'y''^2 + F' = 0. \qquad (I)$$

II. $A' \neq 0$ o $C' = 0$, ma $E' \neq 0$. Lasciando $\xi' = -D'/A'$, $\eta' = 0$, ovvero $y'' = y'$, otteniamo l'equazione

$$A'x''^2 + 2E'y' + F' = 0,$$

O

$$A'x''^2 + 2E'\left(y' + \frac{F'}{2E'}\right) = 0.$$

Quindi effettuando una traslazione parallela lungo l'asse Oy′ di una quantità $\eta'' = -F'/2E'$, troviamo che $y' = y'' - F'/2E'$, cioè $y' + F'/2E' = y''$ in modo da ottenere l'equazione

$$A'x''^2 + 2E'y'' = 0. \qquad (II)$$

Se abbiamo $A' = 0$, $C' \neq 0$, $D' \neq 0$, possiamo semplicemente scambiare i ruoli di x e y e ottenere lo stesso risultato.

III. $A' \neq 0$, $C' = 0$, $E' = 0$. Riprendendo $\xi' = -D'/A'$, $\eta' = 0$, otteniamo l'equazione

$$A'x''^2 + F' = 0. \qquad \text{(III)}$$

Se abbiamo $A' = 0$, $C' \neq 0$, $D' = 0$, possiamo nuovamente scambiare i ruoli di x e y.

Abbiamo ora considerato tutte le possibilità, tenendo conto del fatto che A' e C' non possono essere contemporaneamente zero, poiché in tal caso il grado dell'equazione si ridurrebbe, e abbiamo visto che con le nostre trasformazioni di coordinate questo grado non cambia.

Pertanto, con la scelta opportuna delle coordinate rettangolari, ogni equazione di secondo grado può essere ricondotta ad una delle tre equazioni cosiddette “ridotte” (I), (II), (III).

Sia l’equazione della forma (I) (in questo caso A' e C' non sono uguali a zero). Se $F' \neq 0$, scrivendo l’equazione (I) come

$$\frac{x''^2}{-F'/A'} + \frac{y''^2}{-F'/C'} - 1 = 0,$$

arriviamo, a seconda dei segni di A', C', F', ad una delle equazioni 1, 2 o 4. Se il denominatore di x''^2 è negativo e quello di y''^2 è positivo, allora dobbiamo anche invertire gli assi $O''x''$ e $O''y''$.

Se $F' = 0$, l'equazione (I) può essere scritta nella forma

$$\frac{x''^2}{1/A'} + \frac{y''^2}{1/C'} = 0,$$

e arriviamo alle equazioni 3 o 5.

Se l'equazione ha la forma (II) (in questo caso A' e E' sono entrambi diversi da zero), allora possiamo scriverla come

$$x''^2 + \frac{2E'}{A'} y'' = 0,$$

e denotando $-E'/A'$ con p e scambiando i nomi degli assi $O''x''$ e $O''y''$ si ottiene l'equazione 6.

Infine, se abbiamo un'equazione della forma (III) (dove $A' \neq 0$), può essere riscritta come $x''^2 + F'/A' = 0$ e si ottiene una delle equazioni 7, 8 o 9.

Questo importante teorema sulla possibilità di ridurre ogni equazione di 2° grado ad una delle 9 forme canoniche fu già esaminato in dettaglio da Eulero. Le argomentazioni nel libro di Eulero differiscono solo nella forma da quelle appena esposte.

§ 9. Rappresentazione di forze, velocità e accelerazioni mediante terne di numeri; teoria dei vettori

Dopo Eulero un passo importante fu compiuto da Lagrange. Nella sua “Meccanica analitica”, pubblicata nel 1788, Lagrange aritmetizzava le forze, le velocità e le accelerazioni nello stesso modo in cui Cartesio aritmetizzava i punti. Questa idea che Lagrange sviluppò nel suo libro prese successivamente la forma della cosiddetta teoria dei vettori e si rivelò di grande aiuto in fisica, meccanica e tecnologia.

Coordinate rettangolari nello spazio. Notiamo innanzitutto che né Cartesio né Newton svilupparono la geometria analitica dello spazio. Ciò venne fatto più tardi, nella prima metà del XVIII secolo, da Laguerre e Clairaut. Per specificare un punto *M* nello spazio selezionarono tre assi tra loro perpendicolari Ox, Oy e Oz e considerarono (figura 39) i valori numerici delle distanze del punto *M* dai piani Oyz, Oxz e *Oxy*, presi con i segni corrispondenti, la cosiddetta ascissa x, ordinata y e altezza z del punto *M*.

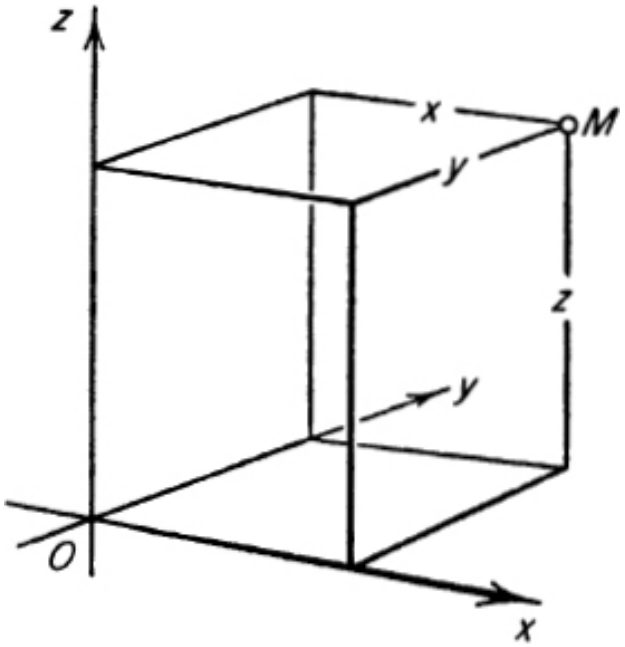


Figura 39.

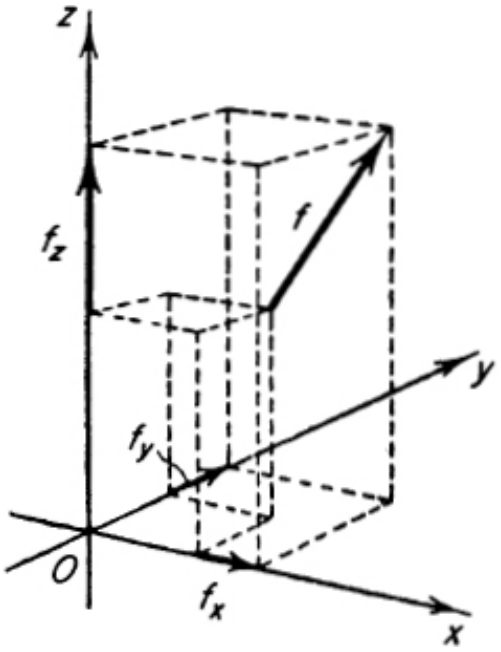


Figura 40.

Aritmetizzazione di forze, velocità e accelerazioni, introdotta da Lagrange. Consideriamo (figura 40) una forza *f* che può essere rappresentata in unità convenzionali da un segmento con una freccia, avente una lunghezza e una direzione specifica. Lagrange sottolinea che questa forza *f* può essere scomposta in tre componenti f_x, f_y e f_z nella direzione dei corrispondenti assi Ox, Oy e Oz; tali componenti, in quanto segmenti diretti sugli assi, possono essere dati semplicemente da numeri, positivi o negativi a seconda che la componente sia diretta nella direzione positiva o opposta dell'asse. Possiamo quindi considerare, ad esempio, la forza (2, 3, 4) oppure la forza (1, -2, 5), ecc. Nella composizione delle forze secondo la legge del parallelogramma, come si può facilmente dimostrare (lo mostreremo più avanti), bisogna sommare le loro componenti corrispondenti. Ad esempio, la somma delle forze date è la forza

$$(2 + 1, 3 — 2, 4 + 5) = (3, 1, 9).$$

Lo stesso può essere fatto per velocità e accelerazioni. In ogni problema di meccanica, tutte le equazioni che collegano forze, velocità e accelerazioni possono anche essere scritte come equazioni che collegano le loro componenti, cioè collegano semplicemente numeri; allora l'equazione meccanica sarà necessariamente scritta sotto forma di tre equazioni: la prima per gli *x*, la seconda per gli *y* e la terza per gli *z*.

Ma fu solo un centinaio d'anni dopo i tempi di Lagrange che matematici e fisici, soprattutto sotto l'influenza della teoria dell'elettricità in via di sviluppo, iniziarono a considerare su larga scala la teoria generale di tali segmenti, aventi una lunghezza e una direzione determinate. Tali segmenti erano chiamati vettori.

La teoria dei vettori ha un grande significato in meccanica, fisica e tecnologia, e il suo lato algebrico, la cosiddetta algebra dei vettori (in contrasto con l'analisi vettoriale) appare subito come una parte costitutiva essenziale della geometria analitica.

Algebra dei vettori. Qualsiasi segmento diretto (sia che rappresenti una forza, una velocità, un'accelerazione o qualche altra entità), cioè un segmento avente una data lunghezza e una direzione definita, è chiamato vettore. Due vettori si dicono uguali se hanno la stessa lunghezza e la stessa direzione; cioè nel concetto stesso di “vettore” si tiene conto solo della sua lunghezza e della sua direzione. È possibile sommare vettori. Siano dati i vettori a, b, ···, d. Disponiamo il vettore a da un punto, poi dal suo punto finale disegniamo il vettore b, ecc. Otteniamo un cosiddetto poligono vettoriale ab ··· d (figura 41). Il vettore m il cui punto iniziale coincide con il punto iniziale del primo vettore a di questo poligono e il cui punto finale coincide con il punto finale dell'ultimo vettore d, è chiamato somma di questi vettori

$$\mathbf{m} = \mathbf{a} + \mathbf{b} + \cdots + \mathbf{d}. \qquad (11)$$

È facile mostrare che il vettore m non dipende dall'ordine in cui si prendono gli addendi a, b, ···, d.

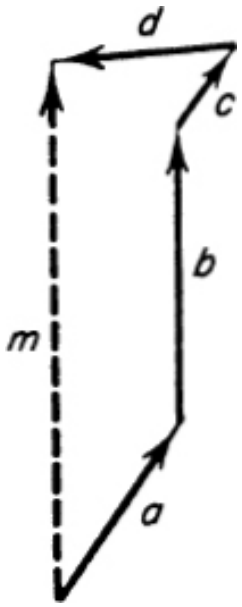


Figura 41.

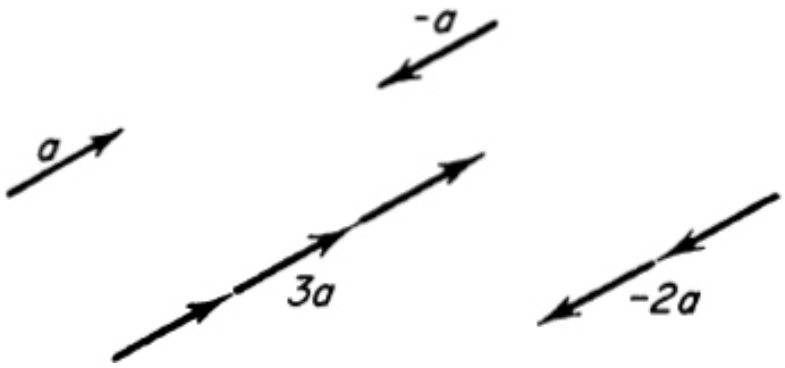


Figura 42.

Il vettore uguale in lunghezza al vettore a ma opposto in direzione è chiamato vettore opposto ed è indicato con -a.

La sottrazione del vettore a è definita come addizione del suo vettore opposto.

Nel calcolo vettoriale i numeri reali ordinari sono comunemente chiamati scalari. Dati un vettore a (figura 42) e uno scalare λ, allora per prodotto del vettore a con lo scalare (numero) λ, cioè λa, si intende il vettore la cui lunghezza è uguale al prodotto della lunghezza $|a|$ del vettore a e il valore assoluto $|\lambda|$ del numero λ, e la cui direzione è la stessa di a se $\lambda > 0$ e opposta se $\lambda < 0$.

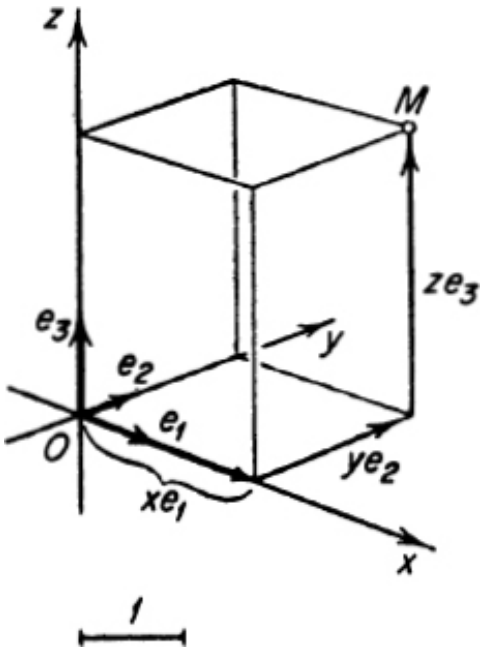


Figura 43.

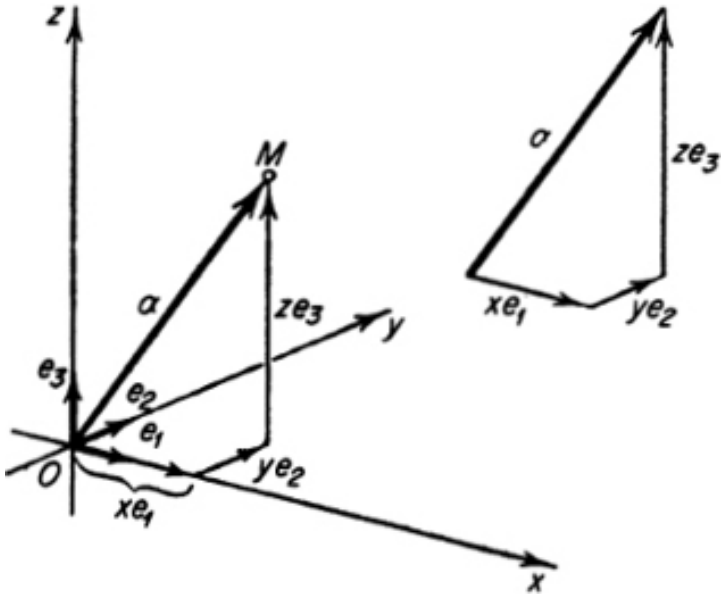


Figura 44.

Consideriamo un sistema di coordinate cartesiane rettangolari *Oxyz* e i vettori e_1, e_2, e_3, aventi lunghezza pari all'unità e direzioni coincidenti rispettivamente con le direzioni positive degli assi Ox, Oy, Oz. È ovvio che un dato punto *M* (figura 43) dello spazio può essere raggiunto dall'origine *O* percorrendo un certo numero di "volte" (un "numero di volte" intero, frazionario o irrazionale, positivo o negativo) il vettore e_1, poi un certo numero di "volte" il vettore e_2, ed infine un certo numero di "volte" il vettore e_3 . È chiaro che i numeri *x*, *y*, *z* che indicano quante "volte" è necessario percorrere i vettori e_1, e_2, e_3, sono semplicemente le coordinate cartesiane del punto *M*.

Sia dato un vettore a; se facciamo spostare un punto dal punto iniziale di a al suo punto finale e scomponiamo questo movimento in movimenti paralleli agli assi Ox, Oy e Oz, e se è quindi necessario spostare il punto di una distanza xe_1 parallela all'asse Ox, attraverso ye_2 parallela all'asse Oy e attraverso ze_3 parallelo all'asse Oz, quindi

$$\mathbf{a} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 . \qquad (12)$$

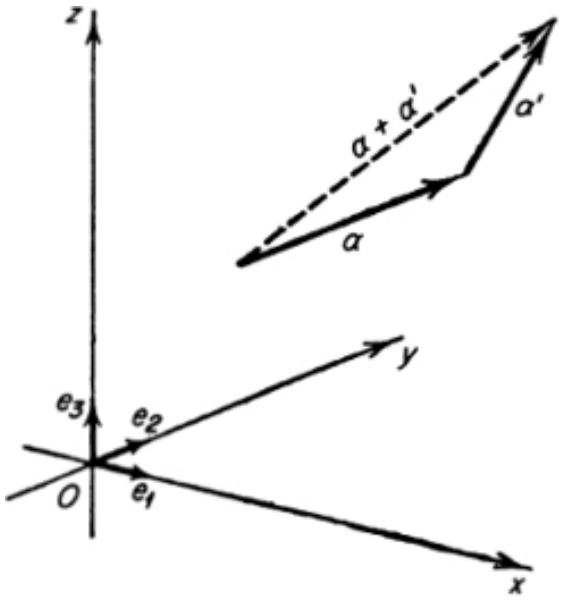


Figura 45.

I numeri *x*, *y*, *z* sono chiamati coordinate del vettore a. Queste sono ovviamente solo le coordinate del punto finale *M* di questo vettore, se il suo punto iniziale si trova nell'origine *O* del sistema di coordinate (figura 44). Da ciò segue subito che sommando i vettori bisogna sommare le loro coordinate corrispondenti, mentre sottraendole devono essere sottratte. Se il primo vettore ci "porta" lungo l'asse Ox di una distanza xe_1, e il secondo di $x'e_1$, allora chiaramente la loro somma ci "porta" lungo l'asse Ox di una distanza $(x + x')e_1$, ecc. (figura 45). Ne consegue subito anche che moltiplicando un vettore per un numero, le sue coordinate vengono moltiplicate per il numero.

Prodotto scalare e sue proprietà. Dati due vettori a e b, il numero |a| |b| cos ϕ, pari al prodotto delle loro lunghezze per il coseno dell'angolo compreso fra essi, è chiamato il loro prodotto scalare ed è indicato con ab o (ab). Siano *x*, *y*, *z* le coordinate del vettore a e $\bar{x}$, $\bar{y}$, $\bar{z}$ le coordinate del vettore b; allora il prodotto scalare è uguale a

$$\mathbf{ab} = x\bar{x} + y\bar{y} + z\bar{z}, \qquad (13)$$

cioè alla somma dei prodotti delle loro coordinate corrispondenti.

Questo importante risultato può essere dimostrato come segue. Innanzitutto facciamo le seguenti osservazioni:

(1) Se moltiplichiamo uno dei vettori di un prodotto scalare, ad esempio a, per un numero λ, ciò equivale ovviamente a moltiplicare il loro prodotto scalare per lo stesso numero, ovvero:

$$(\lambda\mathbf{a})\mathbf{b} = \lambda(\mathbf{ab}).$$

(2) Il prodotto scalare è distributivo, ovvero se $a = a_1 + a_2$, allora $ab = a_1b + a_2b$.

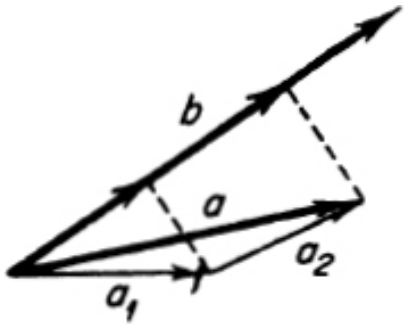


Figura 46.

Infatti, il membro sinistro di questa uguaglianza è pari al prodotto della lunghezza del vettore b per il valore numerico della proiezione del vettore a sull'asse del vettore b (figura 46), e il membro destro è pari al prodotto della lunghezza di b per la somma dei valori numerici delle proiezioni dei vettori a_1 e a_2, sull'asse di b. Ma proj a = proj a_1 + proj a_2, il che dimostra l'uguaglianza.

Consideriamo ora due vettori a e b le cui scomposizioni in termini di vettori e_1, e_2, e_3 sono $\mathbf{a} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$, $\mathbf{b} = \bar{x}\mathbf{e}_1 + \bar{y}\mathbf{e}_2 + \bar{z}\mathbf{e}_3$, cosicché

$$\mathbf{ab} = (x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3)(\bar{x}\mathbf{e}_1 + \bar{y}\mathbf{e}_2 + \bar{z}\mathbf{e}_3).$$

Per la distributività (2) del prodotto scalare, le somme dei vettori tra parentesi possono essere moltiplicate come polinomi, e per (1) i fattori scalari in ciascuno dei termini possono essere presi al di fuori delle parentesi, in modo che

$$\begin{aligned}\mathbf{ab} = {} & x\bar{x}\mathbf{e}_1\mathbf{e}_1 + x\bar{y}\mathbf{e}_1\mathbf{e}_2 + x\bar{z}\mathbf{e}_1\mathbf{e}_3 + y\bar{x}\mathbf{e}_2\mathbf{e}_1 + y\bar{y}\mathbf{e}_2\mathbf{e}_2 + y\bar{z}\mathbf{e}_2\mathbf{e}_3 \\ & + z\bar{x}\mathbf{e}_3\mathbf{e}_1 + z\bar{y}\mathbf{e}_3\mathbf{e}_2 + z\bar{z}\mathbf{e}_3\mathbf{e}_3 .\end{aligned}$$

Ma

$$|\mathbf{e}_1| = |\mathbf{e}_2| = |\mathbf{e}_3| = 1, \quad \cos 0° = 1 \text{ and } \cos 90° = 0.$$

Di conseguenza,

$$\begin{aligned} \mathbf{e}_1\mathbf{e}_1 &= 1, \quad \mathbf{e}_1\mathbf{e}_2 = 0, \quad \mathbf{e}_1\mathbf{e}_3 = 0, \\ \mathbf{e}_2\mathbf{e}_1 &= 0, \quad \mathbf{e}_2\mathbf{e}_2 = 1, \quad \mathbf{e}_2\mathbf{e}_3 = 0, \\ \mathbf{e}_3\mathbf{e}_1 &= 0, \quad \mathbf{e}_3\mathbf{e}_2 = 0, \quad \mathbf{e}_3\mathbf{e}_3 = 1. \end{aligned}$$

Così,

$$\mathbf{ab} = x\bar{x} + y\bar{y} + z\bar{z}. \qquad (14)$$

Osserviamo, in particolare, che se i vettori a e b sono tra loro perpendicolari, allora $\phi = 90°$ e $\cos\phi = 0$. Pertanto l'uguaglianza

$$x\bar{x} + y\bar{y} + z\bar{z} = 0 \qquad (15)$$

serve come condizione facilmente verificabile di perpendicolarità dei vettori a e b.

Angolo tra due direzioni. Consideriamo una direzione caratterizzata dai suoi angoli α, β, γ con gli assi coordinati. Tracciamo la linea

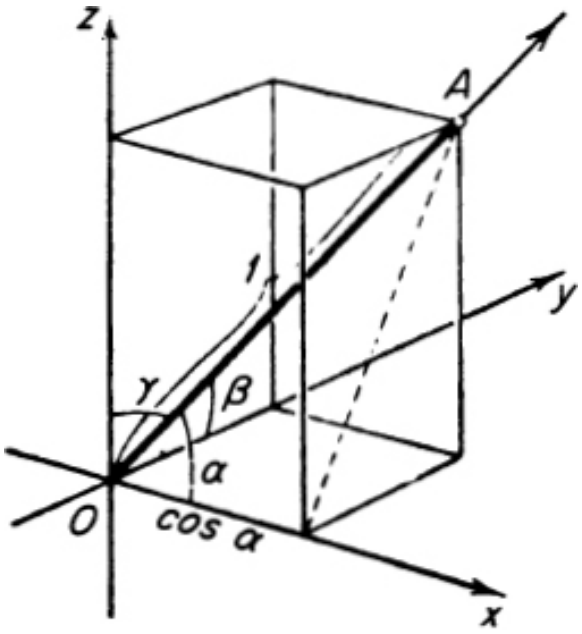


Figura 47.

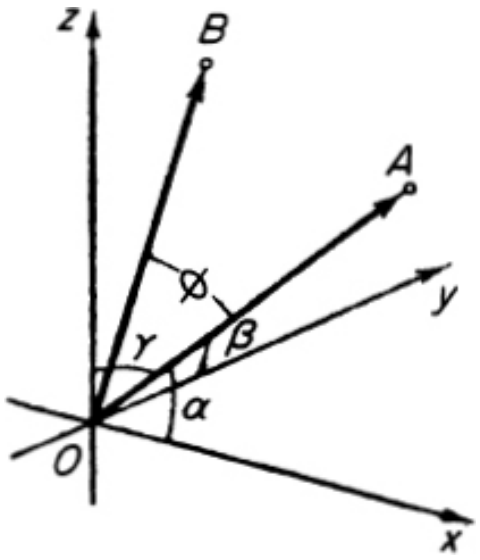


Figura 48.

in questa direzione attraverso l'origine del sistema di coordinate e riportiamo su di essa dall'origine un segmento OA di lunghezza unitaria (figura 47). In questo caso le coordinate del punto A, cioè le coordinate del vettore $\overrightarrow{OA}$ sono esattamente cos α cos β e cos γ. Se abbiamo una seconda direzione data dagli angoli $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$, allora l'analogo vettore $\overrightarrow{OB}$ per questa seconda direzione ha coordinate $\cos\bar{\alpha}$, $\cos\bar{\beta}$, $\cos\bar{\gamma}$ (figura 48). Sia ϕ l'angolo tra questi vettori; allora il loro prodotto scalare è pari a $1 \cdot 1 \cos\phi$ da cui ricaviamo

$$\cos\phi = \cos\alpha\cos\bar{\alpha} + \cos\beta\cos\bar{\beta} + \cos\gamma\cos\bar{\gamma}. \qquad (16)$$

Questa è la formula molto importante per il coseno dell'angolo compreso tra due direzioni.

§ 10. Geometria analitica nello spazio; equazioni di superfici e curve

Se viene data un'equazione $z = f(x, y)$ e se x e y sono considerati come l'ascissa e l'ordinata e z la quota di un punto, allora questa equazione stessa rappresenta una superficie P, che può essere ottenuta erigendo perpendicolari di lunghezza z nei punti (x, y) del piano Oxy. Il luogo dei punti finali di queste perpendicolari dà la superficie P rappresentata da questa equazione. Se l'equazione che collega x, y e z non è già risolta rispetto a z, allora può essere risolta per z e successivamente possiamo costruire la superficie P. In generale, in geometria analitica si dice superficie rappresentata dall'equazione la totalità di tutti quei punti dello spazio le cui coordinate x, y, z soddisfano una data equazione (figura 49) in tre variabili x, y, z.

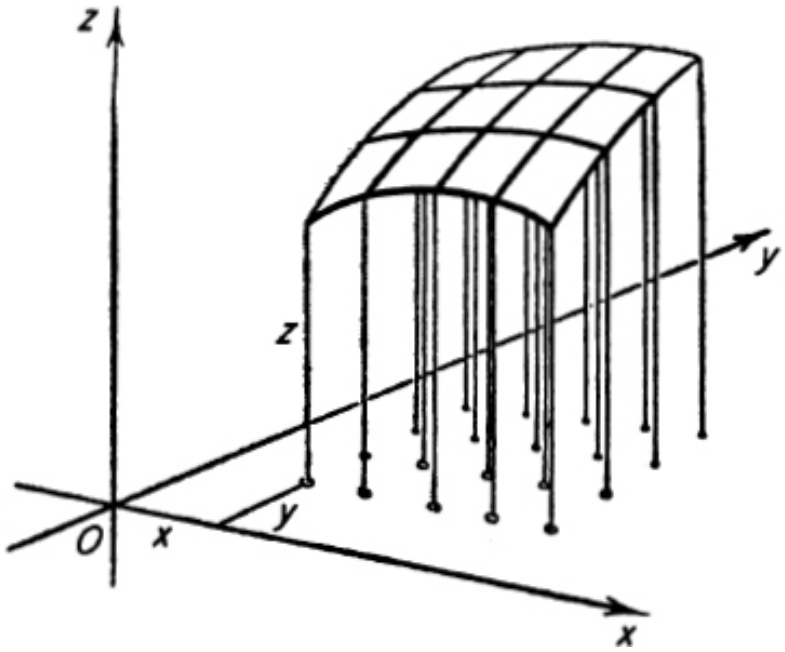


Figura 49.

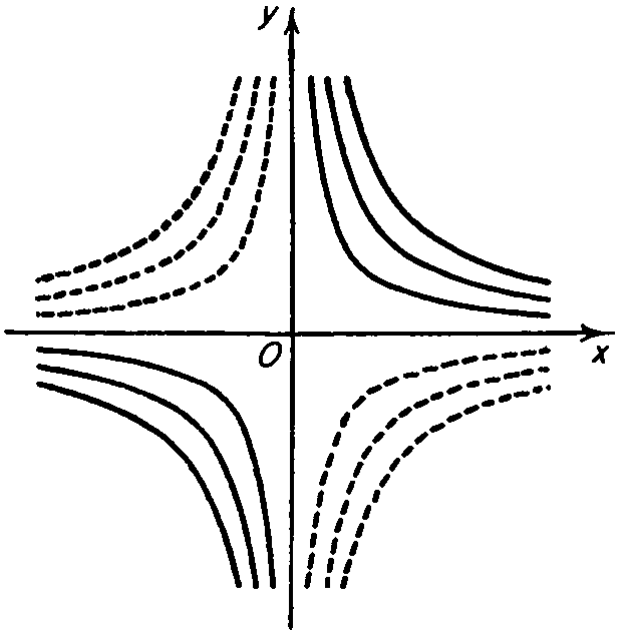


Figura 50.

Una funzione di due variabili $f(x, y)$, come già sottolineato nel Capitolo II, può rappresentare non solo una superficie P, ma anche il suo sistema di curve di livello, cioè curve nel piano Oxy su ciascuna delle quali la funzione $f(x, y)$ ha un valore costante. Questo sistema di curve non è altro chiaramente che la mappa topografica della superficie P sul piano Oxy.

Esempio. L'equazione $xy = z$ dà, ad esempio, le curve di livello: ⋯, $xy = -3$, $xy = -2$, $xy = -1$, $xy = 0$, $xy = 1$; $xy = 2$, $xy = 3$, ⋯. Sono tutte iperboli (figura 50) tranne $xy = 0$, che rappresenta i due assi coordinati. Ciò che si ottiene è chiaramente una superficie a sella (figura 51) (il cosiddetto paraboloide iperbolico).

Per definire una curva nello spazio, possiamo dare le equazioni di due superfici qualsiasi P e Q che si intersecano lungo la curva. Ad esempio, il sistema

$$\begin{aligned} xy &= z, \\ x^2 + y^2 &= 1 \end{aligned}$$

dà una curva spaziale (figura 52). L'equazione $xy = z$ determina il precedente paraboloide iperbolico e l'equazione $x^2 + y^2 = 1$ determina un cilindro circolare di raggio unitario, il cui asse è l'asse Oz. Il sistema di equazioni definisce conseguentemente la curva di intersezione del paraboloide con il cilindro, che è rappresentata in figura 52.

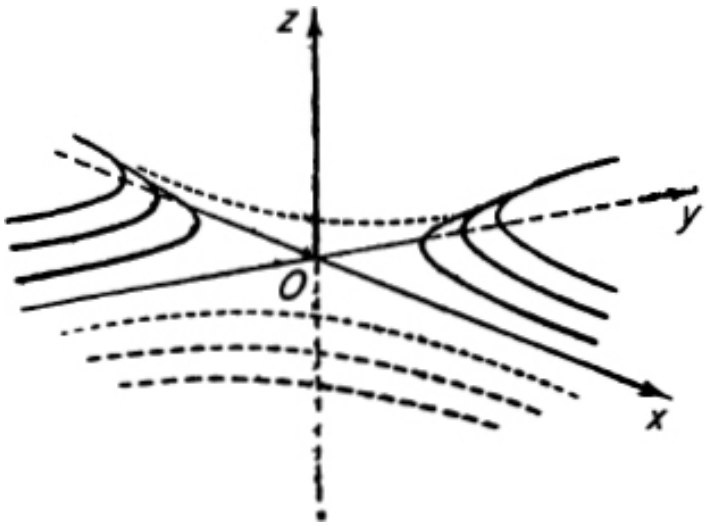


Figura 51.

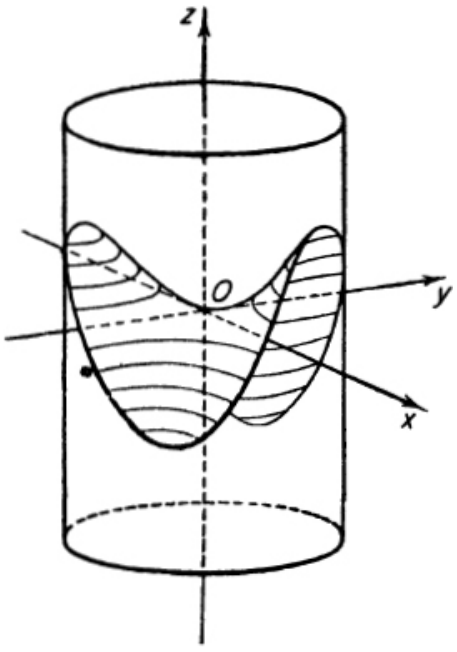


Figura 52.

Se in questo sistema si sceglie arbitrariamente una delle incognite, ad esempio x, e poi si risolve il sistema rispetto a y e z, otterremo le coordinate x, y, z dei vari punti della curva.

Equazione del piano ed equazioni della retta. Si può dimostrare che ogni equazione di primo grado con tre variabili

$$Ax + By + Cz + D = 0$$

rappresenta un piano e viceversa. Da quanto detto risulta chiaro che una retta può essere data da un sistema di due di tali equazioni:

$$\begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0, \end{aligned}$$

cioè, come la curva di intersezione di due piani.

L'equazione generale di secondo grado in tre variabili e le sue 17 forme canoniche. Equazione di secondo grado in tre variabili

$$\begin{aligned} A_1x^2 + A_2y^2 + A_3z^2 + 2B_1yz + 2B_2xz + 2B_3xy \\ + 2C_1x + 2C_2y + 2C_3z + D = 0, \quad (17) \end{aligned}$$

contiene 10 termini. Analogamente a quanto fatto in precedenza per un'equazione a due variabili, si può dimostrare che mediante un'opportuna rotazione del sistema di coordinate dato attorno all'origine, l'equazione (17) può essere ridotta alla forma

$$A_1'x'^2 + A_2'y'^2 + A_3'z'^2 + 2C_1'x' + 2C_2'y' + 2C_3'z' + D = 0, \qquad (18)$$

cioè in modo da eliminare i termini con prodotti delle variabili. In questo caso, però, la dimostrazione della possibilità di una tale semplificazione dell'equazione è notevolmente più difficile che nel caso del piano. La difficoltà della dimostrazione nasce dal fatto che nel piano una rotazione attorno ad un punto è data da un angolo ϕ, da noi scelto opportunamente, mentre nello spazio la rotazione di un corpo attorno ad un punto fisso è data da tre angoli indipendenti (angoli di Eulero) ϕ, θ, ψ e in modo abbastanza complicato. Quindi l'equazione deve essere ripulita dai termini con prodotti di variabili in modo indiretto (vedi Capitolo XVI sulla teoria della riduzione mediante trasformazioni ortogonali di una forma quadratica a una somma di quadrati). Successivamente, come nel caso del piano, si effettua una traslazione parallela degli assi e si semplifica l'equazione, dopodiché l'equazione (18) assume infine una delle seguenti forme canoniche:

1. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} - 1 = 0$  Ellissoide

2. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} + 1 = 0$  Ellissoide immaginario

3. $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} - 1 = 0$  Iperboloide a una falda

4. $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} + 1 = 0$  Iperboloide a due falde

5. $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$  Cono di secondo ordine

6. $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$  Cono immaginario di secondo ordine

7. $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 2cz = 0$  Paraboloide ellittico

8. $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 2cz = 0$  Paraboloide iperbolico

9. $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$  Cilindro ellittico

10. $\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$  Cilindro ellittico immaginario

11. $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0$  Coppia di piani immaginari incidenti

12. $\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$  Cilindro iperbolico

13. $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0$  Coppia di piani incidenti

14. $y^2 - 2px = 0$  Cilindro parabolico

15. $x^2 - a^2 = 0$  Coppia di piani paralleli

16. $x^2 + a^2 = 0$  Coppia di piani paralleli immaginari

17. $x^2 = 0$  Coppia di piani coincidenti

Le ultime nove equazioni canoniche 9-17 non contengono termini in z e rappresentano esattamente le equazioni canoniche delle curve del secondo ordine nel piano *Oxy*. Nello spazio queste equazioni rappresentano cilindri, le cui direttrici sono le corrispondenti curve del secondo ordine nel piano *Oxy* e le cui generatrici sono paralleli all'asse Oz. Infatti, se una di queste equazioni è soddisfatta da un punto con coordinate (x_1, y_1, 0), allora sarà soddisfatta anche da qualsiasi punto con coordinate (x_1, y_1, z) qualunque sia z, poiché nell'equazione non ci sono comunque termini con z.

Tra le equazioni 1-8 come si può facilmente vedere, l'equazione 2 non è soddisfatta da nessun punto con x, y, z reali e l'equazione 6 è soddisfatta solo da uno di questi punti (0, 0, 0), cioè l'origine. Resta quindi da studiare solo le sei equazioni 1, 3, 4, 5, 7, 8.

Ellissoide. Confrontiamo le superfici rappresentate dalle equazioni $x^2/a^2 + y^2/b^2 + z^2/c^2$-1 = 0 e $x^2 + y^2 + z^2$-1 = 0. La seconda di queste è ovviamente l'equazione di una sfera *C* con centro nell'origine e di raggio unitario, poiché $x^2 + y^2 + z^2$ è il quadrato della distanza dal punto (x, y, z) all'origine *O*. Se (x, y, z) è un punto che giace sulla sfera, cioè soddisfa la seconda equazione, allora (ax, by, cz) è un punto le cui coordinate soddisfano la prima equazione. La superficie rappresentata dalla prima equazione si ottiene quindi dalla sfera *C* se si sostituiscono tutte le ascisse x dei punti della sfera con ax, y con by e z con cz, ovvero se la sfera *C* è allungata uniformemente a partire dai piani Oyz, Oxz e *Oxy* con coefficienti di dilatazione a, b e c, rispettivamente. Questa superficie è chiamata ellissoide (figura 53).

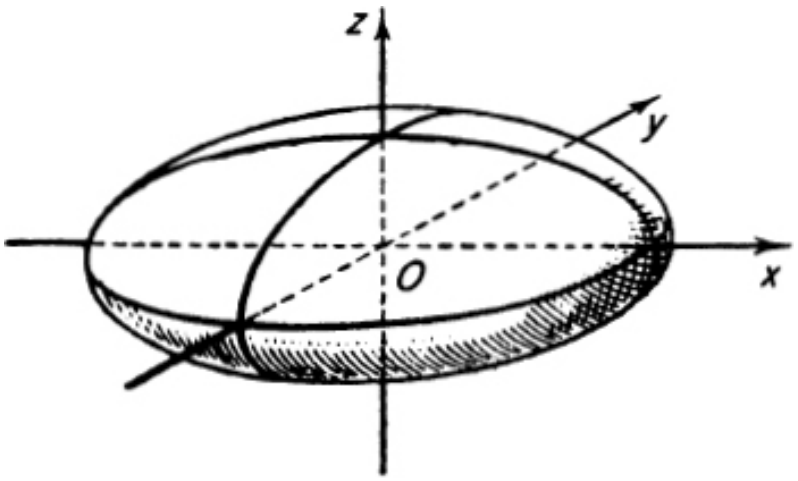


Figura 53.

Iperboloidi e cono del secondo ordine. Consideriamo le equazioni 3, 4 e 5, ovvero l'equazione della forma

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = \delta, \qquad (19)$$

dove $\delta = 1$, -1 o 0. Confrontiamolo con l'equazione

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} - \frac{z^2}{c^2} = \delta, \qquad (20)$$

in cui il denominatore di y^2 è anche a^2 e non b^2, come nell'equazione (19). Come prima, osserviamo che la superficie (19) è ottenuta dalla superficie (20) per espansione dal piano Oxz con coefficiente b/a.

Vediamo ora quale superficie è rappresentata dalla (20). Prendiamo un piano $z = h$ perpendicolare all'asse Oz ed esaminiamo la sua intersezione con la superficie (20). Sostituendo $z = h$ nell'equazione (20), otteniamo

$$x^2 + y^2 = a^2\left(\delta + \frac{h^2}{c^2}\right).$$

Se $\delta + h^2/c^2$ è positivo, allora questa equazione insieme a $z = h$ dà un cerchio, che giace nel piano $z = h$ con centro sull'asse Oz. Se $\delta + h^2/c^2$ è negativo, cosa che può verificarsi solo con $\delta = -1$ e h^2 piccolo, allora il piano $z = h$ non interseca affatto la superficie (20), poiché la somma dei quadrati $x^2 + y^2$ non può essere un numero negativo.

L'intera superficie (20) è quindi costituita da cerchi che giacciono su piani perpendicolari all'asse Oz e aventi i centri sull'asse Oz. Ma in questo caso la superficie (20) è una superficie di rivoluzione attorno all'asse Oz.

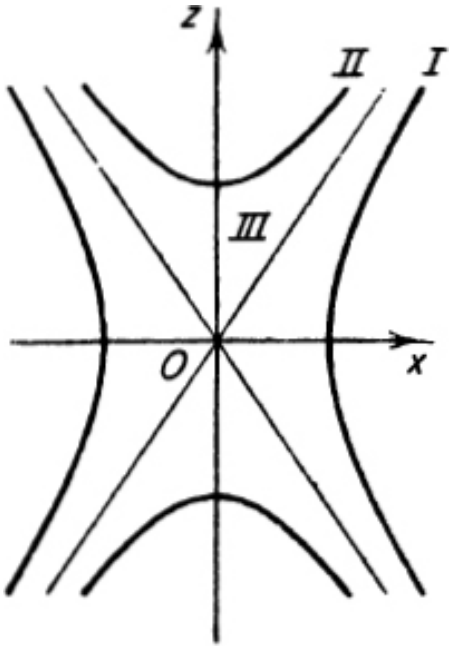


Figura 54.

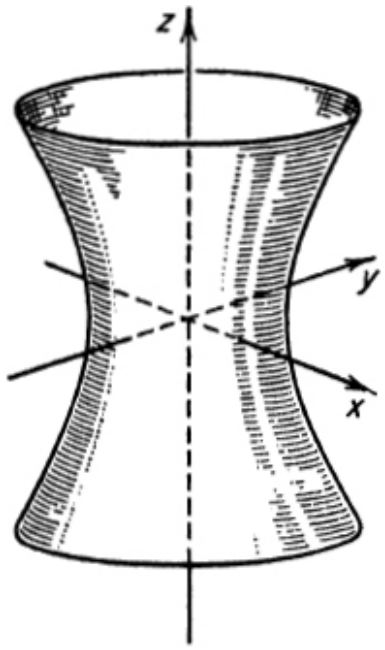


Figura 55.

Se la intersechiamo con un piano passante per l'asse Oz, otteniamo il suo “meridiano”, cioè una curva, giacente in un piano passante per l'asse, dalla rivoluzione della quale si genera la superficie.

Se intersechiamo la superficie (20) con il piano delle coordinate Oxz, cioè il piano $y = 0$ (figura 54), sostituendo $y = 0$ nell'equazione (20), otteniamo l'equazione del meridiano $x^2/a^2 - z^2/c^2 = \delta$. Nel caso $\delta = 1$ questa è l'iperbole *I*, per $\delta = -1$ è l'iperbole II, e per $\delta = 0$ la coppia di rette intersecanti III. Questi, ruotando attorno all'asse *z*, producono rispettivamente un cosiddetto iperboloide di rivoluzione a una falda (figura 55), un iperboloide di rivoluzione a due falde (figura 56) e un cono circolare retto (figura 57).

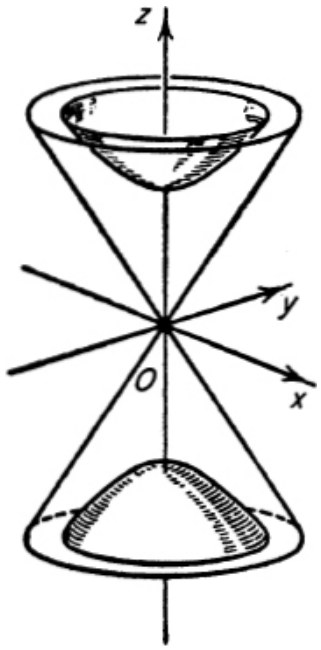


Figura 56.

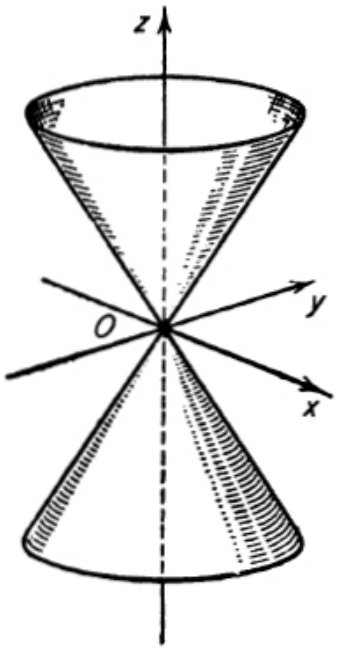


Figura 57.

L'iperboloide generale a una falda, l'iperboloide a due falde e il cono del secondo ordine 3, 4 e 5 si ottengono da queste superfici di rivoluzione mediante un'espansione dal piano Oxz con coefficiente b/a.

Paraboloidi. Rimangono solo le equazioni 7 e 8. Confrontiamo il primo di questi $x^2/a^2 + y^2/b^2 = 2cz$ con l'equazione

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} = 2cz,$$

che indagheremo nello stesso modo di prima. Rappresenta una superficie ottenuta facendo ruotare attorno all'asse Oz la parabola $x^2 = 2a^2cz$, cioè il cosiddetto paraboloide di rivoluzione (figura 58) discusso in precedenza a proposito degli specchi parabolici. Il paraboloide ellittico generale 7 si ottiene dal paraboloide di rivoluzione mediante un'espansione dal piano Oxz.

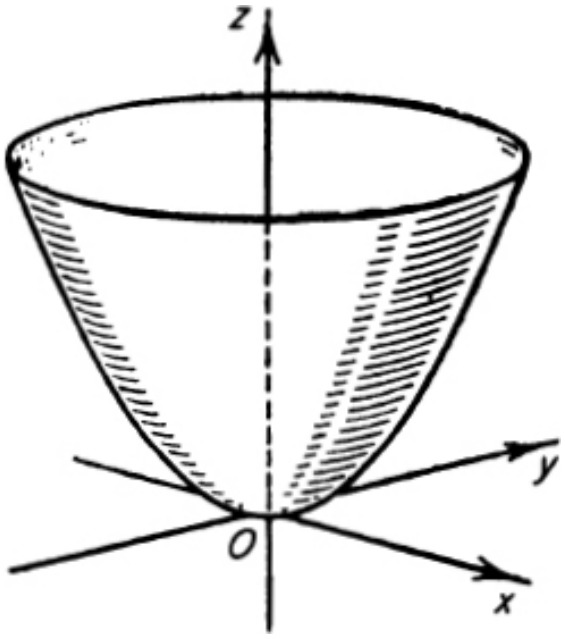


Figura 58.

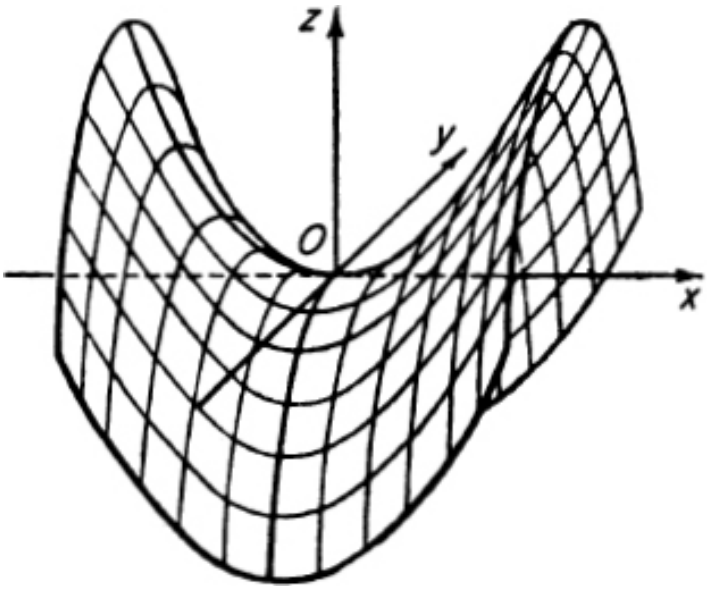


Figura 59.

La superficie 8 deve essere studiata in modo diverso, cioè esaminando le sue intersezioni con i piani $z = h$, che sono iperboli. La mappa di contorno della superficie 8 è rappresentata in figura 50; in una diversa posizione degli assi coordinati abbiamo considerato questa superficie in figura 51. Essa ha la forma di una sella, come illustrato in figura 59 ed è chiamata paraboloide iperbolico. Le sue intersezioni con i piani paralleli al piano Oxz risultano essere parabole identiche. Lo stesso risultato si ottiene mediante intersezioni con piani paralleli al piano Oyz.

Generatori rettilinei di un iperboloide a una falda. È un fatto molto curioso e per nulla ovvio che l'iperboloide a una falda e il paraboloide iperbolico si possano ottenere, proprio come il cono e il cilindro, dal moto di una retta. Nel caso dell'iperboloide, è sufficiente dimostrare

questo fatto per un iperboloide di rivoluzione a una falda $x^2/a^2 + y^2/a^2 - z^2/c^2 = 1$, poiché l'iperboloide generale a una falda è ottenuto da un'espansione uniforme dal piano Oxz e sotto tale espansione qualsiasi retta diventerà una retta. Intersechiamo l'iperboloide di rivoluzione con il piano $y = a$ parallelo al piano Oxz. Sostituendo $y = a$ otteniamo[E1]

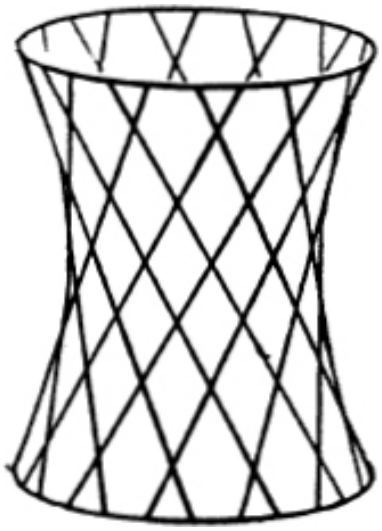


Figura 60.



Figura 61.

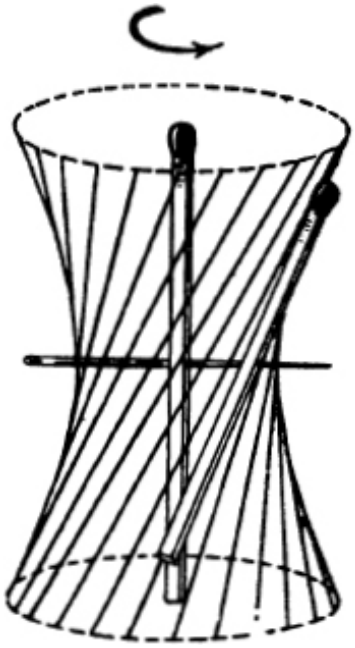


Figura 62.

$$\frac{x^2}{a^2} + \frac{a^2}{a^2} - \frac{z^2}{c^2} = 1 \quad \text{or} \quad \frac{x^2}{a^2} - \frac{z^2}{c^2} = 0.$$

Ma questa equazione insieme a $y = a$ dà nel piano $y = a$ una coppia di rette che si intersecano: $x/a - z/c = 0$ e $x/a + z/c = 0$.

Abbiamo quindi già scoperto che sull'iperboloide giace una coppia di rette che si intersecano. Se ora ruotiamo l'iperboloide attorno all'asse Oz, allora ciascuna di queste linee traccia ovviamente l'intero iperboloide (figura 60).

È facile dimostrare che: (1) due rette arbitrarie di una e della stessa famiglia di linee così ottenute non giacciono nello stesso piano (cioè sono rette sghembe), (2) qualsiasi linea di una di queste famiglie interseca tutte le linee dell'altra famiglia (eccetto la sua opposta, che è parallela ad essa), e (3) tre linee di una e la stessa famiglia non sono parallele a uno stesso piano.

Con due fiammiferi e un ago è facile ottenere una rappresentazione dell'iperboloide di rivoluzione a una falda. Perforiamo uno dei fiammiferi al centro con l'ago e sulla punta appuntita dell'ago fissiamo l'altro fiammifero parallelamente al primo. Se poi ruotiamo l'intero apparecchio attorno al primo fiammifero come asse, il secondo fiammifero traccerà la superficie di un cilindro (figura 61). Ma se il secondo fiammifero non è parallelo al primo, allora durante una rivoluzione traccerà la superficie di un iperboloide di rivoluzione a una falda, come si può facilmente visualizzare se la rotazione è rapida (figura 62).

Riassunto dello studio dell'equazione di secondo grado. Sebbene l'equazione generale di secondo grado a tre variabili possa rappresentare essenzialmente 17 superfici diverse, non è difficile ricordarle. Le ultime nove sono cilindri sopra le nove possibili curve del secondo ordine, mentre le prime otto sono divise in quattro coppie: due ellissoidi (reale e

immaginario), due iperboloidi (a una e a due falde), due coni del secondo ordine (reale e immaginario) e due paraboloidi (ellittici e iperbolici). Tutte queste superfici svolgono un ruolo essenziale nella meccanica, nella fisica e nella tecnologia (ellissoide di inerzia, ellissoide di elasticità, iperboloide nella trasformazione di Lorentz in fisica, paraboloide di rivoluzione per specchi parabolici, ecc.).

§ 11. Trasformazioni affini e ortogonali

Il successivo passo importante nello sviluppo della geometria analitica fu l'introduzione in essa, e nella geometria in generale, della teoria delle trasformazioni. Qui sarà necessario spiegare la questione in qualche dettaglio.

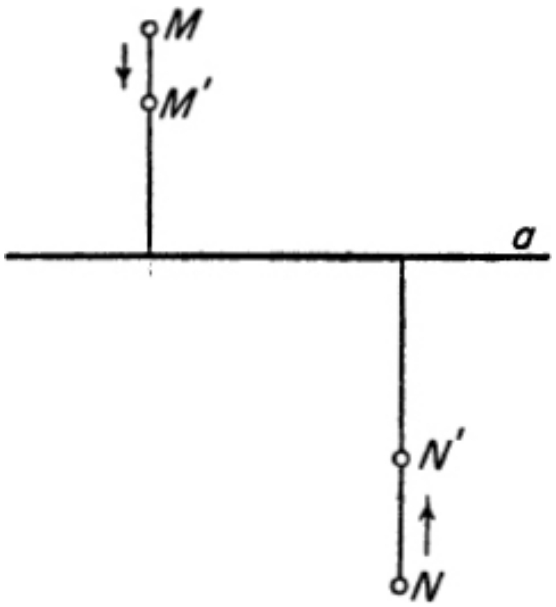


Figura 63.

“Contrazione” del piano verso una linea. Consideriamo una delle trasformazioni più semplici del piano, ovvero la “contrazione” uniforme verso una linea di coefficiente k. Nel piano sia data una linea a e un coefficiente positivo k, ad esempio $k = 2/3$. Tutti i punti della linea a sono fissi e ogni punto M che non giace su questa linea viene inviato nel punto M' in modo tale che M' si trovi sullo stesso lato della linea di M sulla perpendicolare da M a a ad una distanza da a uguale a 2/3 della distanza da M a a. Se il coefficiente k, come in questo caso, è minore dell'unità, allora abbiamo una contrazione propria del piano rispetto alla retta; ma se k è maggiore dell'unità, abbiamo un'espansione del piano dalla retta, ma per comodità in questo ed altri casi parleremo di “contrazione”, salvo che la parola “contrazione” sarà messa tra virgolette.

Il punto o la figura da trasformare si chiama preimmagine e quella in cui viene inviata è la sua immagine. Il punto M', ad esempio, è l'immagine del punto M (figura 63).

Mostriamo che sotto una “contrazione” uniforme di un piano su una linea, qualsiasi linea del piano si trasforma in una linea. Supponiamo infatti che il piano sia “contratto” su una linea *a* giacente in esso con coefficiente di “contrazione” *k*. Sia *b* una qualsiasi linea del piano, *O* il punto in cui interseca la linea *a*, *B* un altro punto arbitrario di *b* e BA la perpendicolare alla linea *a* dal punto *B* (figura 64). Nella “contrazione” il punto *B* va al punto *B′* su questa perpendicolare tale che $B'A = k \cdot BA$. Pertanto, la tangente dell'angolo B′OA sarà uguale a $AB'/OA = k \cdot AB/OA$, cioè sarà uguale a *k* volte la tangente dell'angolo che la linea *b* forma con la linea *a*, cioè per tutti i punti *B′* in cui si incontrano diversi punti della linea *b* si trasformeranno, sarà lo stesso. Tutti i punti *B′* giacciono quindi su una stessa retta, che passa per il punto *O* e forma con la retta *a* un angolo con questa tangente.

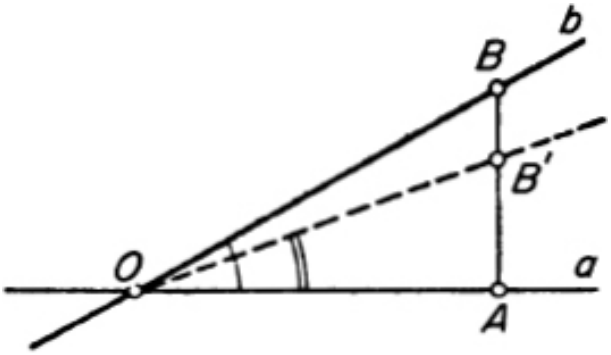


Figura 64.

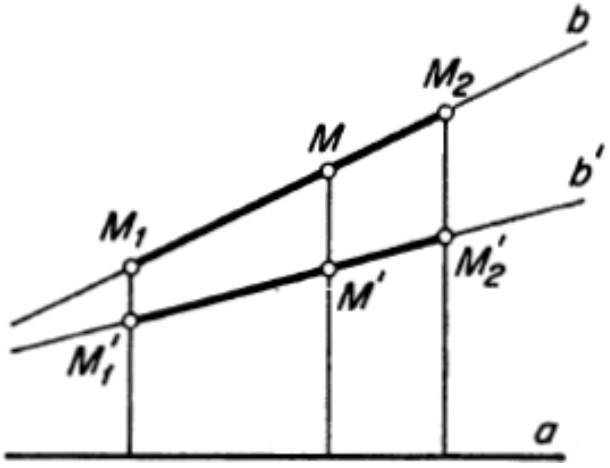


Figura 65.

In “contrazione” le rette parallele rimangono parallele. Infatti, se le tangenti degli angoli che le linee *b* e *c* formano con la linea *a* sono le stesse, allora le tangenti di quegli angoli che le immagini *b′* e *c′* formano con *a* differiscono da loro solo per un fattore *k*, cioè sono ancora uguali tra loro, il che significa che anche le linee *b′* e *c′* sono parallele tra loro.

Qualsiasi segmento rettilineo del piano sotto “contrazione” rispetto a una linea viene contratto (o espanso) in modo uniforme (sebbene a vari gradi per segmenti di varie direzioni). Quando parliamo qui di contrazione “uniforme”, intendiamo che il punto medio del segmento rimane il punto medio, il terzo rimane il terzo, ecc., cioè il segmento si restringe uniformemente su tutta la sua lunghezza. Infatti, in qualunque rapporto il punto *M* divide il segmento M_1M_2, la

sua immagine M' dividerà $M'_1M'_2$, nello stesso rapporto, poiché le rette parallele (in questo caso perpendicolari alla linea a) tagliano le linee che le intersecano (in questo caso b e b') in parti proporzionali (figura 65).

L'ellisse come risultato della “contrazione” di un cerchio. Consideriamo una circonferenza con centro nell'origine e raggio a. Per il teorema di Pitagora la sua equazione è $x^2 + \bar{y}^2 = a^2$, dove abbiamo scritto $\bar{y}$ invece di y, poiché y ci servirà in seguito. Vediamo in cosa si contrae questo cerchio se “contraiamo” il piano sull'asse Ox con coefficiente b/a (figura 66). Dopo questa “contrazione” i valori x di tutti i punti rimangono gli stessi, ma i valori $\bar{y}$ diventano uguali a $y = \bar{y}\,(b/a)$, cioè $\bar{y} = (a/b)\,y$. Sostituendo $\bar{y}$ nell'equazione del cerchio sopra, avremo:

$$x^2 + \frac{a^2}{b^2}\,y^2 = a^2 \quad \text{or} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

come l'equazione, nello stesso sistema di coordinate, della curva ottenuta dal cerchio dato per contrazione rispetto all'asse Ox. Come vediamo, otteniamo un'ellisse. Abbiamo così dimostrato che un'ellisse è il risultato della “contrazione” di un cerchio.

Dal fatto che un'ellisse è una “contrazione” di un cerchio derivano direttamente molte proprietà dell'ellisse. Ad esempio, la già citata proprietà dei diametri, cioè che se di un'ellisse sono date secanti parallele, allora i loro punti medi giacciono su una retta (vedi figura 12), può essere mostrata nel modo seguente. Eseguiamo l'espansione inversa dell'ellisse nel cerchio. Sotto questa espansione le corde parallele dell'ellisse vanno nelle corde parallele del cerchio, ed i loro punti medi nei punti medi di queste corde. Ma i punti medi delle corde parallele di un cerchio giacciono su un diametro, cioè su una retta, così che anche i punti medi delle corde parallele dell'ellisse giacciono su una retta. Esse giacciono cioè su quella linea che si ottiene dal diametro del cerchio sotto la “contrazione” che manda il cerchio nell'ellisse.

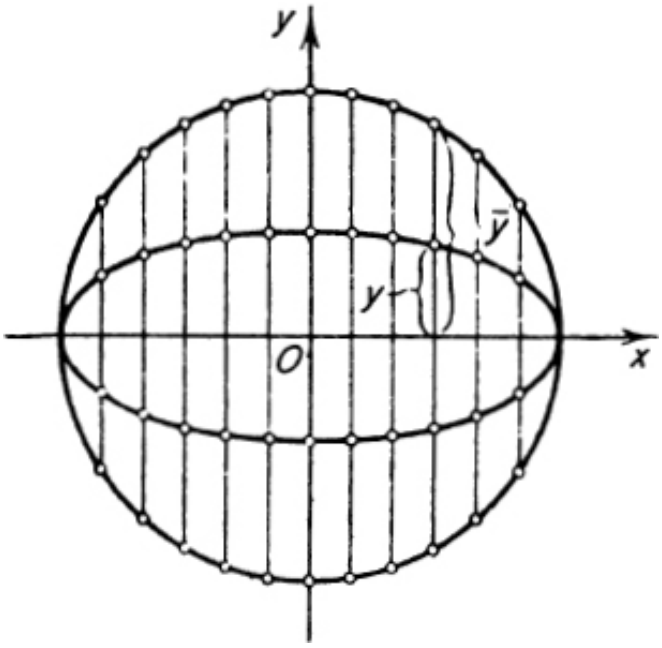


Figura 66.

Ecco un'altra applicazione della teoria della "contrazione". Poiché qualsiasi striscia verticale del cerchio sotto la sua contrazione verso l'asse Ox non cambia la sua larghezza e la sua lunghezza è moltiplicata per b/a, l'area di questa striscia dopo la contrazione è uguale alla sua area iniziale moltiplicata per b/a, e poiché l'area del cerchio è uguale a πa^2, l'area dell'ellisse corrispondente è uguale a $\pi a^2(b/a) = \pi ab$.

Esempio della soluzione di un problema più complicato. Sia data un'ellisse e sia richiesto di trovare il triangolo con l'area più piccola circoscritta a questa ellisse. Per prima cosa risolviamo il problema per un cerchio. Mostriamo che nel caso di un cerchio, questo è un triangolo equilatero. Sia infatti non equilatero il triangolo circoscritto; cioè, il più piccolo dei suoi angoli (indicato con B) è inferiore a 60° e il più grande $C > 60°$. Se poi, senza variare l'angolo A, spostiamo il lato BC nella posizione B_0C_0 (figura 67) spostando il vertice B verso A fino ad uno degli angoli B_0 oppure C_0 diventa uguale a 60°, otteniamo un triangolo circoscritto AB_0C_0

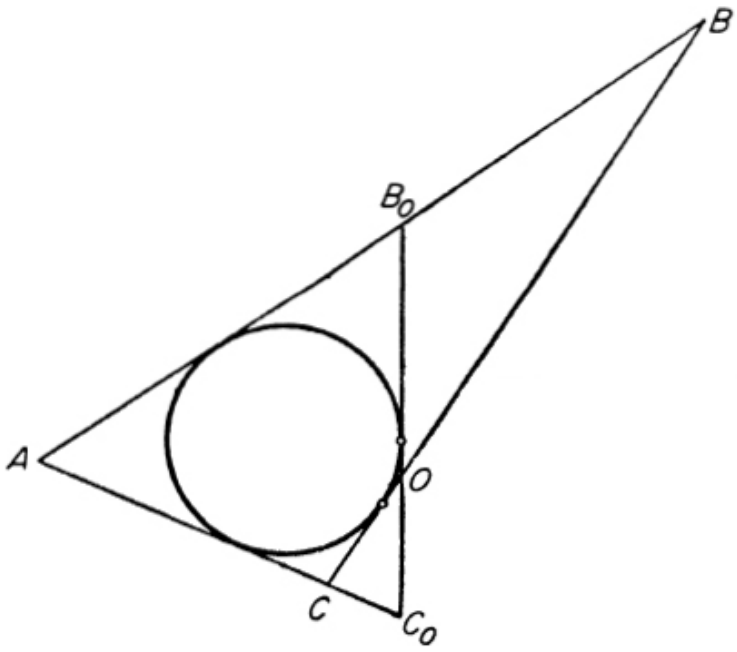


Figura 67.

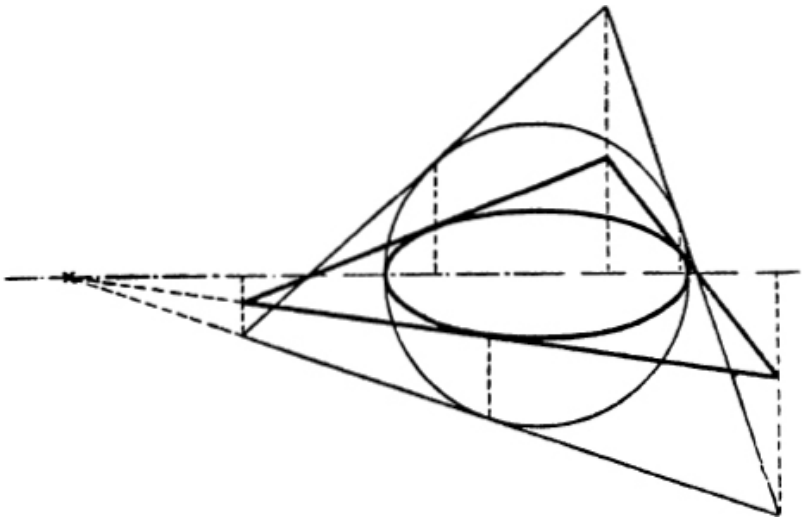


Figura 68.

con area minore, poiché qui[4] OC < OB, $OC_0 \leqslant OB_0$ e quindi l'area scartata OBB_0 è maggiore di quella aggiunta OCC_0. Se il triangolo così ottenuto non è equilatero, allora ripetendo il procedimento sopra descritto riduciamo ulteriormente la sua area e arriviamo ad un triangolo equilatero. Quindi, qualsiasi triangolo non equilatero circoscritto a un dato cerchio ha un'area maggiore di quello equilatero.

Torniamo ora all'ellisse. Facciamone un'espansione dall'asse maggiore, riconvertendolo così nel cerchio da cui è stato ottenuto per “contrazione”. Con questa espansione (figura 68): (1) tutti i triangoli circoscritti all'ellisse si trasformano in triangoli circoscritti al cerchio risultante; (2) le aree di tutte le figure, e in particolare di questi triangoli, vengono aumentate nello stesso rapporto. Da ciò vediamo che i triangoli che circoscrivono l'ellisse data con l'area più piccola saranno quelli che verranno convertiti in triangoli equilateri che circoscrivono il cerchio. Esistono infiniti triangoli di questo tipo; ognuno di essi ha il baricentro al centro dell'ellisse ed i punti di tangenza sono al centro dei suoi lati. Uno qualsiasi di questi triangoli può essere facilmente costruito (figura 68), a partire dal cerchio sopra menzionato.

Le “contrazioni” del piano rispetto ad una linea sono solo un caso particolare delle trasformazioni più generali, cosiddette affini, del piano.

Trasformazioni affini generali. Una coppia di vettori e_1, e_2 non allineati, con origine comune *O*, sarà chiamata riferimento cartesiano del piano. Le coordinate di un punto *M* rispetto a questo riferimento Oe_1e_2 sono i numeri *x*, *y* tali che, per raggiungere *M* da *O*, si percorre prima *x* volte il vettore e_1 e poi *y* volte il vettore e_2. Si ottiene così un sistema generale di coordinate cartesiane del piano. Un sistema analogo può essere introdotto nello spazio. Il consueto sistema di coordinate cartesiane ortogonali, finora utilizzato, corrisponde al caso particolare in cui i vettori e_1, e_2 sono perpendicolari e hanno lunghezza uguale all’unità di misura.

Una trasformazione affine generale del piano è quella in cui una data rete di parallelogrammi uguali viene trasformata in un'altra rete arbitraria di parallelogrammi uguali. Più precisamente si tratta di una trasformazione del piano sotto la quale un dato sistema di

coordinate Oe_1e_2 si trasforma in un certo altro sistema (in generale con un'altra "metrica", cioè con lunghezze diverse per i vettori e'_1 e e'_2 e un angolo diverso tra loro) e un punto arbitrario M viene inviato nel punto M' avente le stesse coordinate relative al nuovo riferimento che M aveva rispetto al vecchio (figura 69).

La “contrazione” all'asse Ox con coefficiente k è un caso particolare in cui il riferimento rettangolare Oe_1e_2 passa nel riferimento Oe_1ke_2.

Si può facilmente dimostrare che con una trasformazione affine ogni retta viene trasformata in una retta, le rette parallele vengono trasformate in rette parallele e se un punto divide un segmento in un dato rapporto, la sua immagine divide l'immagine di questo segmento nello stesso rapporto. Inoltre, possiamo dimostrare il notevole teorema secondo cui qualsiasi trasformazione affine del piano può essere ottenuta eseguendo un certo movimento rigido del piano su se stesso, e quindi, in generale, due “contrazioni” con coefficienti diversi k_1 e k_2 a due rette tra loro perpendicolari.

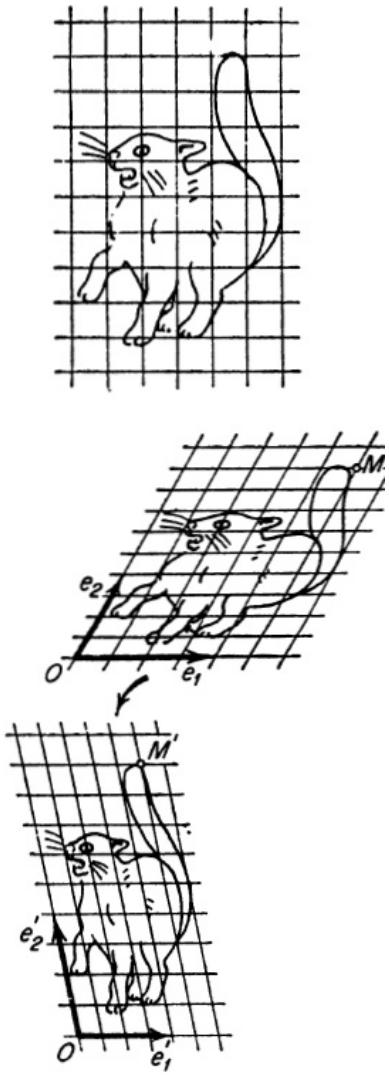


Figura 69.

Per dimostrare questa affermazione consideriamo tutti i raggi di una circonferenza del piano (figura 70). Sia il raggio OA quello che, dopo la trasformazione, risulta essere il più corto, e

venga mappato in $O'A'$. La perpendicolare AB a OA viene quindi trasformata in $A'B'$, che deve essere perpendicolare a $O'A'$, poiché se la perpendicolare $O'C'$ fosse diversa da $O'A'$, allora sarebbe l'immagine dell'obliquo OC, e l'immagine $O'D'$ del raggio OD sarebbe una parte della perpendicolare $O'C'$, cioè più corto dell'obliquo $O'A'$, contrariamente a quanto ipotizzato.

Le linee reciprocamente perpendicolari OA e AB vengono quindi mappate nelle linee reciprocamente perpendicolari $O'A'$ e $A'B'$. Di conseguenza, la rete quadrata costruita su OA e AB si trasforma in una rete di rettangoli uguali (figura 71) e lungo le rette di questa rete quadrata avvengono "contrazioni" uniformi.

In modo del tutto analogo una trasformazione affine generale dello spazio può essere definita come quella in cui un sistema di coordinate spaziali $Oe_1e_2e_3$ viene trasformato in qualche altro sistema $O'e'_1e'_2e'_3$, in generale, con una "metrica" diversa, cioè con segmenti unitari di diversa lunghezza e con diversi angoli tra loro, e un punto M viene inviato nel punto M' avente le stesse coordinate relative al nuovo riferimento di quelle del punto M rispetto al vecchio riferimento.

Tutte le proprietà qui enumerate valgono anche per trasformazioni affini dello spazio, tranne che nell'ultimo teorema ci sarà un movimento rigido dello spazio e poi tre "contrazioni" a tre piani tra loro perpendicolari con determinati coefficienti k_1, k_2, k_3.

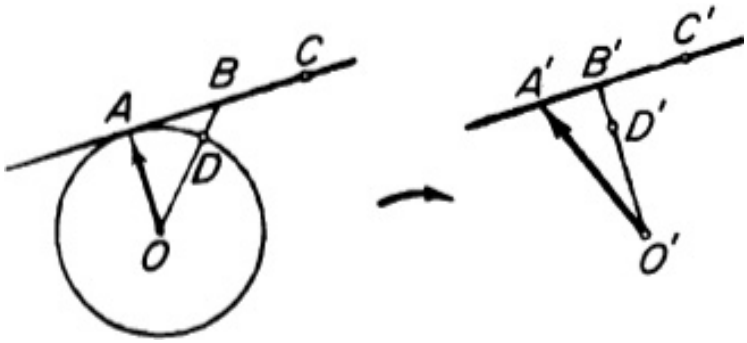


Figura 70.

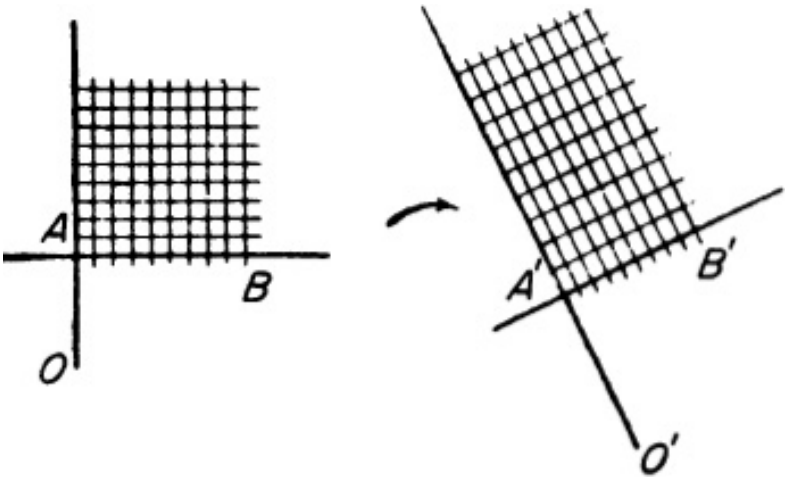


Figura 71.

Applicazioni delle trasformazioni affini. Le applicazioni più importanti delle trasformazioni affini sono:

1. In primo luogo c'è l'applicazione in geometria alla soluzione di problemi riguardanti proprietà affini delle figure, cioè proprietà che si conservano sotto trasformazioni affini. Ne sono esempi il teorema sui diametri di un'ellisse e il problema dei triangoli circoscritti. Per risolvere tali problemi effettuiamo una trasformazione affine della figura in una più semplice, per la quale dimostriamo la proprietà desiderata e poi torniamo alla figura originale.

2. In secondo luogo, vi è l'applicazione in geometria analitica alla classificazione di curve e superfici del secondo ordine. Il punto principale è, come si può dimostrare, che diverse ellissi sono legate tra loro nel senso che l'una può essere ottenuta da un'altra mediante una trasformazione affine (la parola latina affinis significa “correlato”). Inoltre tutte le iperboli sono affini tra loro, e lo sono anche tutte le parabole. Ma non possiamo convertire un'ellisse in una parabola, o un'iperbole in una parabola, mediante una trasformazione affine, cioè non sono affini tra loro. È naturale dividere tutte le curve del secondo ordine in classi affini di curve, affine tra loro. Risulta che la riduzione di un'equazione alla forma canonica dà esattamente questa classificazione; cioè, ci sono nove classi affini di curve del secondo ordine. (Non entreremo nei dettagli sul motivo per cui ellissi immaginarie e coppie di rette parallele immaginarie appartengono a classi affini diverse. Propriamente parlando, né in un caso né nell'altro ci sono curve sul piano. La questione qui riguarda in realtà le proprietà algebriche dell'equazione stessa.)

Allo stesso modo, la classificazione delle superfici del secondo ordine secondo le loro equazioni canoniche in 17 forme è la stessa della classificazione affine.

Diamo un semplice esempio di applicazione della classificazione affine delle superfici del secondo ordine. Mostriamo che se selezioniamo arbitrariamente nello spazio tre linee *a*, *b*, *c* tali che (1) due qualsiasi di esse non giacciono sullo stesso piano (cioè sono sghembe) e (2) non sono tutte parallele allo stesso piano, allora l'insieme di tutte le rette *d* dello spazio, ciascuna delle quali simultaneamente interseca tutte e tre le linee date *a*, *b*, *c* (figura 72) costituisce l'intera superficie di un iperboloide a una falda.

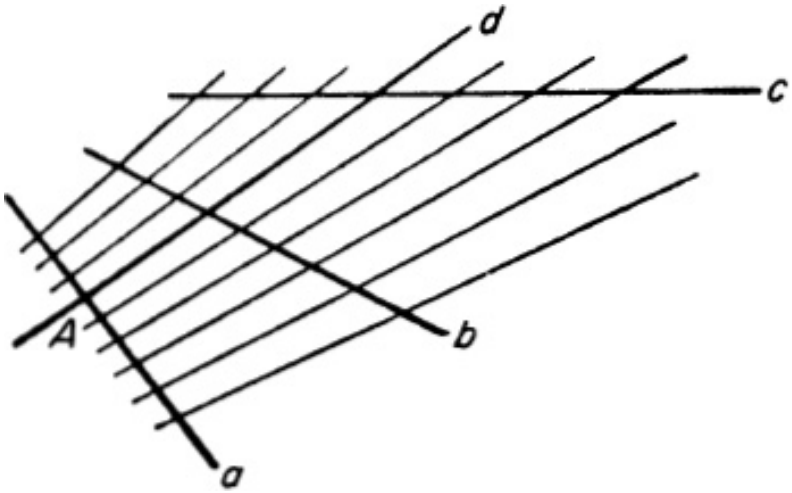


Figura 72.

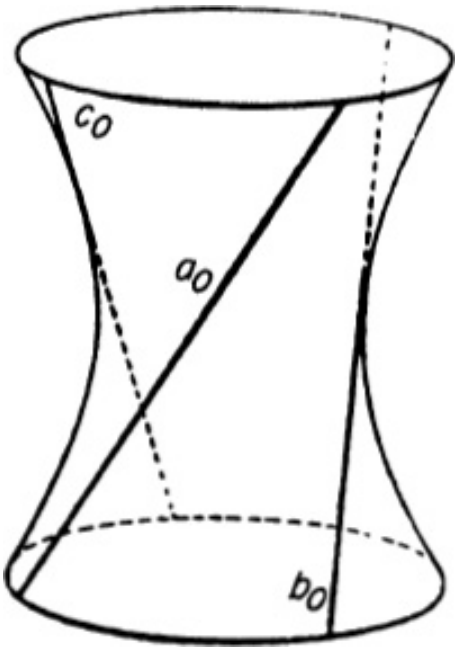


Figura 73.

Spieghiamo più in dettaglio l'insieme di linee *d* di cui stiamo discutendo qui. Attraverso un punto arbitrario *A* della linea *a*, possiamo passare un piano *P* contenente la linea *b* e un piano *Q* contenente la linea *c*. Questi piani *P* e *Q* si intersecano in un'unica linea *d*, che passa per il punto *A* della linea *a* e interseca le linee *b* e *c*. Disegnando tutte queste linee *d* attraverso punti arbitrari della linea *a*, otteniamo l'insieme di tutte quelle linee *d* dello spazio, ciascuna delle quali interseca tutte e tre le linee date *a*, *b* e *c*. Questo insieme di linee determina una superficie. Notiamo che ogni dato iperboloide a una falda può essere ottenuto in questo modo, poiché per le rette *a*, *b* e *c* basta prendere solo tre rette distinte a_0, b_0, c_0 di una famiglia (figura 73) e per le linee *d* tutte le rette dell'altra famiglia. Al contrario, siano date tre rette sghembe arbitrarie dello spazio *a*, *b*, *c*, non tutte parallele allo stesso piano. Allora, come si può vedere, queste linee formano sempre i tre bordi (senza punti in comune) di qualche parallelepipedo (figura 74). Dopo aver costruito tali parallelepipedi per le linee *a*, *b*, *c* e per tre linee a_0, b_0, c_0, di una stessa famiglia di iperboloide arbitrario a una falda, eseguiamo una trasformazione affine dello spazio che invia il parallelepipedo a_0, b_0, c_0, nel parallelepipedo *a*, *b*, *c*; ovviamente questa trasformazione mappa l'iperboloide sulla superficie in questione. Ma secondo la classificazione affine delle superfici del secondo ordine, l'immagine affine di un iperboloide a una falda è ancora un iperboloide a una falda.

3. In terzo luogo, c'è l'applicazione alla teoria delle trasformazioni continue di mezzi continui, ad esempio, nella teoria dell'elasticità, nella teoria dei campi elettrici o magnetici, ecc. Elementi molto piccoli di un dato mezzo continuo si trasformano “quasi” in modo affine. Diciamo: “nel piccolo la trasformazione è lineare” (chiamiamo lineare un'espressione di primo grado, e nel paragrafo successivo vediamo che in geometria analitica le formule delle trasformazioni affini sono di primo grado). Ciò è evidente nella figura 75. Sulle linee del grande reticolo quadrato si nota chiaramente la loro distorsione o “distensione a ventaglio”. Ma per un piccolo pezzo della fittissima rete quadrata, tutto questo si mostra ben poco, e la rete quadrata si trasforma “quasi” in una rete di parallelogrammi uguali. Un riferimento

simile si ottiene anche nello spazio (figura 76). Poiché ogni trasformazione affine dello spazio si riduce a un movimento e a tre “contrazioni” reciprocamente perpendicolari, ne consegue che un elemento di un corpo sottoposto a deformazione elastica si muove prima come un corpo rigido e poi subisce tre “contrazioni” reciprocamente perpendicolari.

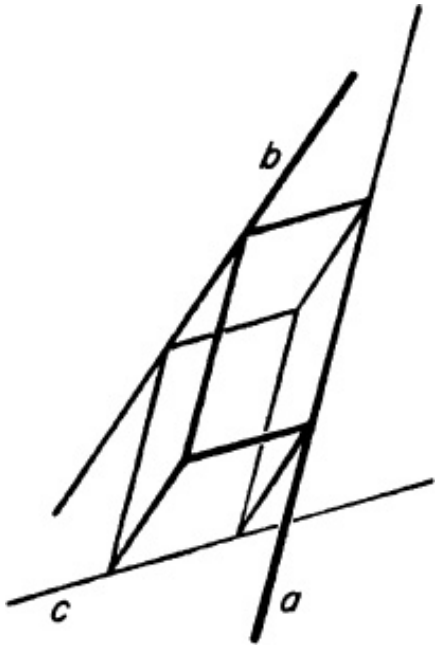


Figura 74.

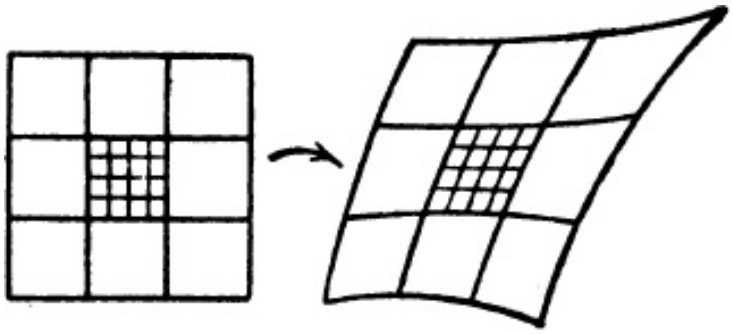


Figura 75.

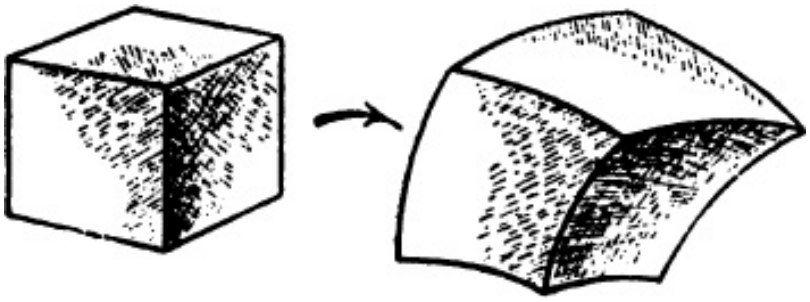


Figura 76.

Formule di trasformazioni affini. Se il riferimento Oe_1e_2 è sottoposto a una trasformazione affine e $O'e'_1e'_2$ ne è l'immagine, mentre le coordinate della nuova origine O' relative al vecchio riferimento sono ξ, η e le coordinate dei vettori e'_1 e e'_2 relativi al vecchio riferimento

sono a_1, a_2 e b_1, b_2, rispettivamente, allora le formule della trasformazione affine, come si vede facilmente dalla figura 77, sono

$$x' = a_1x + b_1y + \xi,$$
$$y' = a_2x + b_2y + \eta$$

nel senso che se x, y sono le coordinate di un qualsiasi punto M relativo al vecchio riferimento Oe_1e_2, allora x', y' dati da queste formule, sono le coordinate relative allo stesso riferimento dell'immagine M' di questo punto.

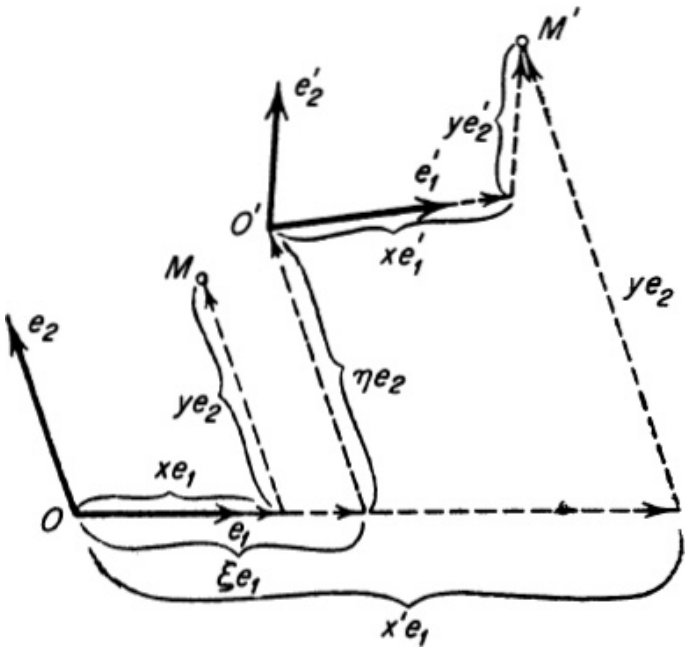


Figura 77.

Infatti, sia Oe_1e_2 un riferimento prima della trasformazione, e $O'e'_1e'_2$ la sua immagine, mentre M sia un punto arbitrario del piano e M' è la sua immagine. Quindi per la stessa definizione di trasformazione affine, se le coordinate del punto M rispetto al riferimento Oe_1e_2 sono x, y, allora le coordinate della sua immagine M' rispetto all'immagine $O'e'_1e'_2$ di questo riferimento sono esattamente le stesse x, y.

Consideriamo ora un vettore m' che unisce l'origine O del vecchio riferimento all'immagine M' del punto M. Quindi $m' = x'e_1 + y'e_2$. Ma questo vettore è uguale a una certa somma vettoriale

$$\boldsymbol{m'} = \xi \boldsymbol{e_1} + \eta \boldsymbol{e_2} + x\boldsymbol{e_1'} + y\boldsymbol{e_2'},$$

e i vettori e'_1 e e'_2 sono

$$\boldsymbol{e_1'} = a_1\boldsymbol{e_1} + a_2\boldsymbol{e_2}, \quad \boldsymbol{e_2'} = b_1\boldsymbol{e_1} + b_2\boldsymbol{e_2}$$

affinché

$$\boldsymbol{m'} = \xi \boldsymbol{e}_1 + \eta \boldsymbol{e}_2 + a_1 x \boldsymbol{e}_1 + a_2 x \boldsymbol{e}_2 + b_1 y \boldsymbol{e}_1 + b_2 y \boldsymbol{e}_2$$

O

$$\boldsymbol{m'} = (a_1 x + b_1 y + \xi)\boldsymbol{e}_1 + (a_2 x + b_2 y + \eta)\boldsymbol{e}_2 .$$

Confrontando questa espressione con la prima espressione per m' otteniamo

$$\left.\begin{aligned} x' &= a_1 x + b_1 y + \xi, \\ y' &= a_2 x + b_2 y + \eta. \end{aligned}\right\} \qquad (21)$$

Il determinante

$$\Delta = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1,$$

come si può vedere, non è zero ed è uguale al rapporto tra l'area del parallelogramma costruito sui vettori del nuovo sistema e l'area dello stesso parallelogramma costruito sui vettori del vecchio sistema.

Formule analoghe si ottengono per lo spazio

$$\left.\begin{aligned} x' &= a_1 x + b_1 y + c_1 z + \xi, \\ y' &= a_2 x + b_2 y + c_2 z + \eta, \\ z' &= a_3 x + b_3 y + c_3 z + \zeta, \end{aligned}\right\} \qquad (22)$$

dove (ξ, η, ζ) sono le coordinate dell'origine O' del riferimento trasformato $O'e'_1e'_2e'_3$ e (a_1, a_2, a_3), (b_1, b_2, b_3), (c_1, c_2, c_3) sono le coordinate dei suoi vettori e'_1, e'_2, e'_3 relative al vecchio riferimento $Oe_1e_2e_3$.

Il determinante[5]

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1$$

non è zero ed è uguale al rapporto tra il volume del parallelepipedo formato dai vettori del nuovo riferimento e il volume del parallelepipedo formato dai vettori del vecchio riferimento.

Trasformazioni ortogonali. I movimenti rigidi del piano su se stesso o tali movimenti più una riflessione attorno a una linea giacente nel piano, sono chiamati trasformazioni ortogonali del piano, e i movimenti rigidi dello spazio, o tali movimenti più una riflessione dello spazio attorno a uno dei suoi piani, sono chiamati trasformazioni ortogonali dello spazio. È chiaro che le trasformazioni ortogonali sono trasformazioni affini sotto le quali la “metrica” del riferimento non cambia, poiché subisce solo un movimento rigido, oppure tale movimento più una riflessione.

Investigheremo le trasformazioni ortogonali mediante coordinate rettangolari, ovvero quando i vettori del sistema originario sono tra loro perpendicolari e hanno lunghezze pari all'unità di misura. Dopo una trasformazione ortogonale i vettori del sistema di riferimento rimangono tra loro perpendicolari, cioè il loro prodotto scalare rimane uguale a zero e le loro lunghezze rimangono uguali all'unità. Pertanto (vedi formula (14), questo capitolo) nel caso del piano, abbiamo

$$a_1b_1 + a_2b_2 = 0, \quad a_1^2 + a_2^2 = 1, \quad b_1^2 + b_2^2 = 1, \qquad (21')$$

e nel caso dello spazio

$$\begin{aligned} &a_1b_1 + a_2b_2 + a_3b_3 = 0, \quad a_1^2 + a_2^2 + a_3^2 = 1, \\ &a_1c_1 + a_2c_2 + a_3c_3 = 0, \quad b_1^2 + b_2^2 + b_3^2 = 1, \qquad (22') \\ &b_1c_1 + b_2c_2 + b_3c_3 = 0, \quad c_1^2 + c_2^2 + c_3^2 = 1. \end{aligned}$$

Quindi, se il sistema di riferimento iniziale è considerato rettangolare, allora le formule (21) danno una trasformazione ortogonale se e solo se le condizioni (21′) di ortogonalità sono soddisfatte, e le formule (22) danno una trasformazione ortogonale dello spazio se le condizioni (22′) di ortogonalità sono soddisfatte. Si può dimostrare che se $\Delta > 0$ abbiamo un movimento rigido, e se $\Delta < 0$ abbiamo un movimento rigido più una riflessione.

§ 12. Teoria degli invarianti

Il concetto di invariante.[6] Invarianti di un'equazione di secondo grado a due variabili. Nella seconda metà del secolo scorso venne introdotto ancora un altro importante concetto, quello di invariante.

Consideriamo, ad esempio, un polinomio di secondo grado in due variabili

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F. \qquad (23)$$

Se consideriamo x, y come coordinate rettangolari ed effettuiamo una trasformazione in nuovi assi rettangolari, quindi dopo aver sostituito x, y nella (23) con le loro espressioni in termini delle nuove coordinate x', y', rimuovendo le parentesi e riducendo termini simili, otteniamo un nuovo polinomio trasformato con coefficienti diversi

$$A'x'^2 + 2B'x'y' + C'y'^2 + 2D'x' + 2E'y' + F'. \qquad (24)$$

Risulta che esistono espressioni formate dai coefficienti che sotto questa trasformazione non cambiano il loro valore numerico, sebbene i coefficienti stessi cambino. Tale espressione in A', B', C', D', E', F' ha esattamente lo stesso valore numerico di quando è formata con A, B, C, D, E, F.

Espressioni di questo tipo sono chiamate invarianti del polinomio (23) rispetto al gruppo di trasformazioni ortogonali (cioè relative alle trasformazioni da un insieme di coordinate rettangolari x, y a qualsiasi altra coordinata rettangolare x', y').

Si può dimostrare che i seguenti sono invarianti di questo tipo:

$$I_1 = A + C,$$
$$I_2 = \begin{vmatrix} A & B \\ B & C \end{vmatrix} = AC - B^2,$$
$$I_3 = \begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} = ACF + 2BDE - AE^2 - CD^2 - FB^2,$$

cioè:

$$A + C = A' + C', \quad AC - B^2 = A'C' - B'^2,$$
$$ACF + 2BDE - AE^2 - CD^2 - FB^2 = A'C'F' + 2B'D'E' - A'E'^2 - C'D'^2 - F'B'^2.$$

È possibile dimostrare l'importante teorema secondo cui qualsiasi invariante ortogonale del polinomio (23) può essere espresso in termini di questi tre invarianti fondamentali.

Se uguagliamo il polinomio (23) a zero, otteniamo un'equazione di una curva del secondo ordine. Qualsiasi quantità, connessa con questa curva ma non con la sua posizione nel piano, ovviamente non dipenderà da quali coordinate è scritta la sua equazione, e quindi, espressa in termini di coefficienti, sarà un invariante ortogonale del polinomio (23), e quindi sarà esprimibile in termini dei tre invarianti fondamentali. Inoltre, poiché moltiplicando tutti e sei i coefficienti dell'equazione per un qualsiasi numero t (diverso da zero) la curva rappresentata dall'equazione rimane la stessa, un'espressione di qualsiasi proprietà della curva in termini di I_1, I_2, I_3, deve certamente essere tale che se i numeri A, B, C, D, E, F in esso contenuti vengono moltiplicati per t, il fattore t si semplifica. L'espressione in questione deve essere, come si suol dire, omogenea di grado zero rispetto a A, B, C, D, E, F.

Verifichiamolo con un esempio. Ad esempio, consideriamo l'equazione

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$$

che rappresenta un’ellisse. Poiché l'equazione determina completamente questa ellisse, possiamo calcolare da essa (cioè dai suoi coefficienti) tutte le quantità fondamentali legate all'ellisse. Ad esempio, possiamo calcolare i suoi semiassi a e b, cioè possiamo esprimere i semiassi in termini di coefficienti. Le espressioni per questi semiassi saranno invarianti e quindi esprimibili in termini di I_1, I_2, I_3. Riducendo l'equazione alla forma canonica e qualche calcolo successivo, si ottengono le seguenti espressioni piuttosto complicate per i semiassi in termini di I_1, I_2, I_3:

$$\sqrt{\frac{2\,|\,I_3\,|}{|\,I_2\,|\,|\,I_1 \pm \sqrt{I_1^2 - 4I_2}\,|}},$$

che sono omogenee rispetto a A, B, C, D, E, F.

Da ciò è chiaro che gli invarianti I_1, I_2, I_3, essi stessi, essendo omogenei ma non di grado zero, non hanno significati geometrici diretti; sono entità algebriche.

Si può dimostrare che l'espressione

$$K_1 = \begin{vmatrix} A & D \\ D & F \end{vmatrix} + \begin{vmatrix} C & E \\ E & F \end{vmatrix} = AF - D^2 + CF - E^2$$

può variare per traslazione parallela ma non per pura rotazione degli assi rettangolari dati, ed è quindi detta seminvariante.

Come esempio di applicazione degli invarianti e dei seminvarianti riportiamo la Tabella 1, che se calcoliamo I_1, I_2, I_3 e K_1, ci permette di determinare direttamente dalla sua equazione la classe affine di una curva del secondo ordine.

Nella tabella 1 sono fornite le condizioni necessarie e sufficienti affinché un'equazione di una curva del secondo ordine sia riducibile all'una o all'altra delle nove forme canoniche ($I_1 I_3$ designa il prodotto di I_1 e I_3).

Tabella 1. Classificazione delle coniche

Criterio	Curva	Equazione ridotta	Equazione canonica
$I_2 > 0, I_1 I_3 < 0$	Ellisse		$x^2/a^2 + y^2/b^2 = 1$
$I_2 > 0, I_1 I_3 > 0$	Ellisse immaginaria		$x^2/a^2 + y^2/b^2 = -1$
$I_2 > 0, I_3 = 0$	Punto	$\lambda_1 x^2 + \lambda_2 y^2 + I_3/I_2 = 0$	$x^2/a^2 + y^2/b^2 = 0$
$I_2 < 0, I_3 \neq 0$	Iperbole		$x^2/a^2 - y^2/b^2 = 1$
$I_2 < 0, I_3 = 0$	Coppia di rette incidenti		$x^2/a^2 - y^2/b^2 = 0$
$I_2 = 0, I_3 \neq 0$	Parabola	$I_1 x^2 + 2\sqrt{(-I_3/I_1)}\, y = 0$	$x^2 = 2py$
$I_2 = 0, I_3 = 0, K_1 < 0$	Coppia di rette parallele		$x^2 = a^2$
$I_2 = 0, I_3 = 0, K_1 > 0$	Coppia di rette parallele immaginarie	$I_1 x^2 + K_1/I_1 = 0$	$x^2 = -a^2$
$I_2 = 0, I_3 = 0, K_1 = 0$	Coppia di rette coincidenti		$x^2 = 0$

La settima riga corregge un refuso dell’originale. [E1]

Consideriamo, ad esempio, l'equazione $x^2 - 6xy + 5y^2 - 2x + 4y + 3 = 0$. Abbiamo $A = 1$, $B = -3$, $C = 5$, $D = -1$, $E = 2$, $F = 3$, quindi $I_1 = 6$, $I_2 = -4$, $I_3 = -9$. Sono soddisfatte le condizioni della 4a riga della tabella: $I_2 < 0$, $I_3 \neq 0$, cioè questa è un'iperbole. I suoi semiassi sono uguali a

$$\sqrt{\frac{2 \cdot 9}{4 \cdot |6 \pm \sqrt{36 + 16}|}} \approx 0.57 \quad \text{and} \quad 1.93.$$

I coefficienti delle equazioni ridotte (I), (II) e (III) sono dati in termini di invarianti e seminvarianti come segue:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{I_3}{I_2} = 0, \qquad \text{(I)}$$

$$I_1 x''^2 + 2\sqrt{-\frac{I_3}{I_1}}\, y'' = 0, \qquad \text{(II)}$$

$$I_1 x''^2 + \frac{K_1}{I_1} = 0, \qquad \text{(III)}$$

dove λ_1 e λ_2 sono le radici della cosiddetta equazione quadratica caratteristica

$$\lambda^2 - I_1\lambda + I_2 = 0.$$

Le formule (I-III) permettono di calcolare velocemente i semiassi a e b di un'ellisse e di un'iperbole, il parametro p di una parabola e la distanza $2a$ tra rette parallele. Le formule per i semiassi sono state fornite in precedenza. Il parametro p è uguale a[E1]

$$p = \sqrt{-\frac{I_3}{I_1^3}} \quad \text{and the distance} \quad 2a = 2\sqrt{-\frac{K_1}{I_1^2}}.$$

Una teoria del tutto analoga degli invarianti e dei seminvarianti, con la relativa tabella per la determinazione della classe affine e le formule dei coefficienti delle equazioni ridotte, può essere data per superfici del secondo ordine nello spazio tridimensionale.

È bene precisare che finora abbiamo trattato solo gli invarianti considerati in geometria analitica per curve e superfici del secondo ordine. Il concetto di invariante, tuttavia, ha un significato molto più ampio.

Per invariante di qualche oggetto in studio, relativo ad alcune delle sue trasformazioni, intendiamo qualsiasi quantità numerica, vettoriale, ecc. connessa a questo oggetto che non varia sotto queste trasformazioni. Nel problema precedente l'oggetto è un polinomio di secondo grado a due variabili (cioè, più precisamente, l'insieme dei suoi coefficienti), e le trasformazioni sono quelle del polinomio ottenuto dalla transizione da un sistema di coordinate rettangolari all'altro.

Un altro esempio: l'oggetto è una data massa di un dato gas a una data temperatura. Le trasformazioni sono cambiamenti di volume o di pressione di questa massa di gas.

L'invariante, secondo la legge Boyle-Mariotte, è il prodotto del volume per la pressione. Possiamo parlare di lunghezze di segmenti nello spazio o di dimensioni di angoli come invarianti del gruppo dei moti dello spazio, di rapporti secondo cui un punto divide un segmento, o di rapporti di aree, come invarianti del gruppo delle trasformazioni affini dello spazio, ecc.

Vari invarianti sono particolarmente importanti in fisica.

§ 13. Geometria proiettiva

Proiezioni prospettiche. Gli artisti hanno cominciato molto tempo fa a studiare le leggi della prospettiva. Ciò era necessario perché l'essere umano vede gli oggetti in proiezione prospettica sulla retina dell'occhio, in modo tale che la forma e la posizione reciproca degli oggetti vengono distorte in modo caratteristico. Ad esempio, i pali del telegrafo in lontananza sembrano più piccoli e più vicini, i binari paralleli di una ferrovia sembrano convergere, ecc. Qui non considereremo le prospettive spaziali, cioè le proprietà delle proiezioni prospettiche degli oggetti nello spazio su un piano, ma solo le proprietà delle proiezioni prospettiche di un piano su un piano.

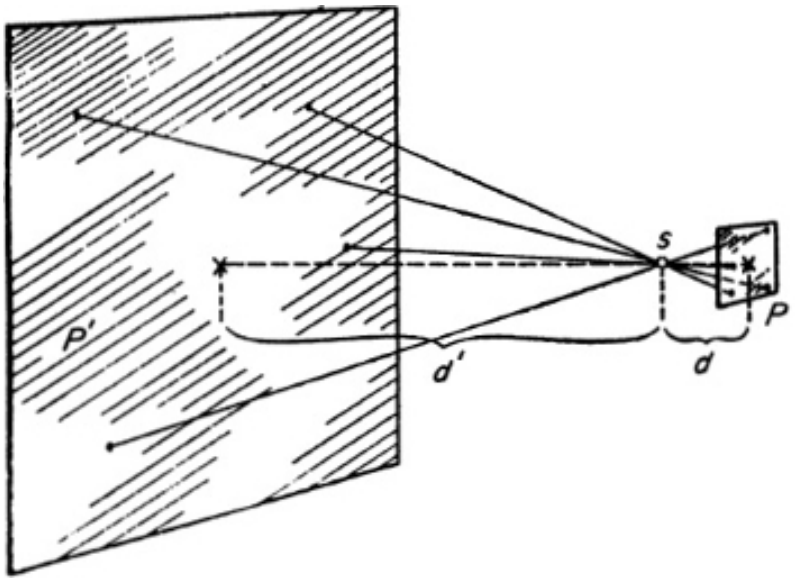


Figura 78.

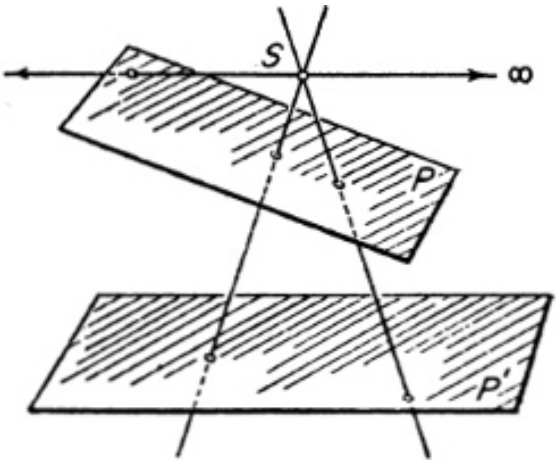


Figura 79.

Consideriamo una fotografia (ad esempio, un fotogramma di una pellicola cinematografica) *P*, uno schermo *P′* e tra loro un obiettivo *S* (figura 78). Quindi, se la fotografia è trasparente ed è illuminata da dietro (se non è trasparente, lasciala essere illuminata dal davanti, cioè dal lato dove si trova l'obiettivo), allora i punti illuminati della fotografia irradiano fasci di luce, che vengono raccolti dall'obiettivo in modo tale da apparire nuovamente sullo schermo *P′* sotto forma di punti. Assumeremo che questa proiezione avvenga come se i punti della fotografia *P* fossero proiettati sullo schermo *P′* su rette passanti per il centro ottico *S* dell'obiettivo.

La situazione sarà molto semplice se i piani *P* e *P′* sono paralleli. In questo caso otterremo ovviamente sul piano *P′* un'immagine non distorta di tutto ciò che si trova sul piano *P*. Questa immagine sarà più piccola o più grande dell'originale a seconda che il rapporto *d′:d*, dove *d* e *d′* sono le distanze dal centro dell'obiettivo ai piani *P* e *P′* rispettivamente, sia più piccolo o più grande di 1.

La situazione sarà notevolmente più difficile se i piani *P* e *P′* non sono paralleli (figura 79). In questo caso, sotto la proiezione attraverso il punto *S* non solo cambia la dimensione della figura ma anche la sua forma viene distorta. Le rette parallele sotto tale proiezione possono diventare convergenti, il rapporto in cui un punto divide un segmento può cambiare, ecc. In generale, alcune delle relazioni che rimangono invarianti sotto una trasformazione affine arbitraria possono cambiare qui.

Questo tipo di proiezione avviene ad esempio nella fotografia aerea. L'aereo oscilla in volo e quindi l'apparecchio fotografico (figura 80a) ad esso rigidamente attaccato non è, in genere, orientato del tutto verticalmente ma al momento dell'esposizione è solitamente in posizione obliqua, cioè otteniamo un'immagine distorta della località (che assumiamo essere piana).

Come possiamo correggere questa immagine? Per questo è necessario studiare le proprietà di proiezione di un piano P su un altro piano Π (in generale i due piani non sono paralleli) mediante rette passanti per un punto *S* che non si trova sul piano *P* né sul piano Π. Tali proiezioni sono chiamate proiezioni prospettiche.

Dimostreremo in seguito il seguente importante teorema.

Teorema. Se abbiamo due proiezioni prospettiche di un piano P sul piano Π tali che sotto entrambe le proiezioni i punti A, *B*, *C*, D di una quadrupla di punti in «posizione generale» sul piano P (cioè una quadrupla in cui nessun gruppo di tre punti è allineato), sono proiettati negli stessi punti A′, *B′*, *C′*, *D′* rispettivamente del piano Π, allora anche tutti i punti del piano P vengono proiettati sotto entrambe le proiezioni negli stessi punti del piano Π.

In altre parole, il risultato di una proiezione prospettica è completamente determinato se si sa in quali punti questa proiezione manda i punti di una quadrupla arbitraria di punti di posizione generale nella figura da proiettare.

Questo è il cosiddetto teorema di unicità della teoria delle trasformazioni proiettive o teorema fondamentale della prospettività del piano.

Applicazione del teorema fondamentale della prospettiva piana nella fotografia aerea. Mostriamo come questo teorema fornisce un metodo adatto per correggere questa immagine in fotografia.

Se al momento dell'esposizione aerea immaginiamo uno schermo orizzontale Π posto alla distanza *h* sotto il centro *S* dell'obiettivo (figura 80a), allora la proiezione su questo schermo attraverso il centro *S* dell'immagine registrata sulla lastra fotografica *P* non sarà ovviamente distorta ma sarà simile alla località orizzontale con una scala *h*: *H*, dove *H* è l'altezza dell'aereo al momento dell'esposizione. Per correggere l'immagine ricevuta sulla fotografia *P* in modo da convertirla in un'immagine non distorta, la trattiamo come segue. La fotografia sviluppata *P* è posta in un apparecchio di proiezione appoggiato su un apposito treppiede sul quale, mediante viti regolabili, l'apparecchio può essere avvicinato o allontanato dallo schermo Π e può essere ruotato in ogni modo.

Allo schermo Π (figura 80b) alleghiamo una mappa topografica della località ricavata da misurazioni sulla superficie terrestre (non una mappa dettagliata, poiché i dettagli che ci interessano devono essere forniti dalla fotografia aerea). Su questa mappa allegata allo schermo Π selezioniamo quattro punti *A′*, *B′*, *C′*, *D′* che si possono ritrovare facilmente anche sulla fotografia (ad esempio un incrocio di strade, un angolo di una casa, ecc.), e nei corrispondenti punti *A*, *B*, *C*, *D* della foto *P* foriamo la pellicola con un ago. Posizioniamo quindi una lampada di proiezione dietro la piastra *P* in una posizione tale che l'immagine venga proiettata sullo schermo Π attraverso una lente *S* dell'apparecchio di supporto. Utilizzando le viti regolabili facciamo in modo che i fasci di luce provenienti dai fori stenopeici cadano sui punti corrispondenti *A′*, *B′*, *C′*, *D′* della mappa attaccata allo schermo. Fatto ciò, sostituiamo la carta topografica con un portalastra con una lastra fotografica e poi, senza modificare la regolazione delle viti, fotografiamo l'immagine proiettata sullo schermo Π della foto *P* scattata dall'aereo.

Per il teorema esposto in precedenza, otteniamo così una mappa corretta (cioè simile alla località) e non distorta della regione fotografata.

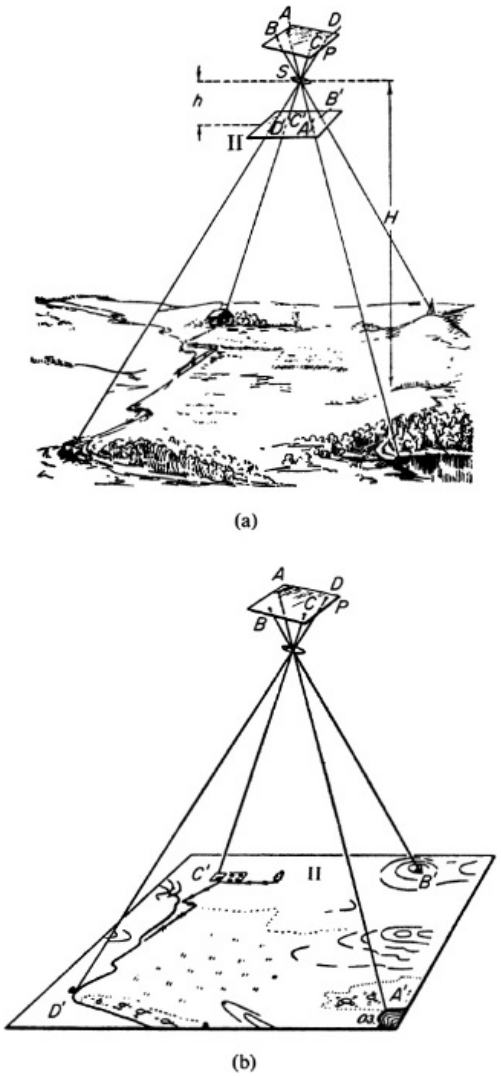


Figura 80.

Passiamo ora alla presentazione della teoria necessaria per dimostrare il teorema fondamentale.

Il piano proiettivo. L'insieme di tutte le linee e i piani dello spazio che passano per un dato punto *S* dello spazio è chiamato fascio proiettante di linee e piani con centro *S*. Se questo fascio è intersecato da un piano *P*, non passante per il centro, allora ad ogni punto del piano *P* corrisponderà una linea del fascio che interseca il piano *P* in questo punto, e ad ogni linea del piano *P* corrisponderà quel piano del fascio che interseca il piano *P* lungo questa linea. Tuttavia, in questo modo non stabiliamo una corrispondenza biunivoca dall'insieme di linee e piani del fascio sull'insieme di punti e linee del piano *P*. Infatti le linee ed i piani del fascio paralleli al piano *P* non corrispondono in questo senso ad alcun punto o linea del piano *P*, poiché non lo intersecano. Tuttavia siamo d'accordo nel dire che queste linee del fascio intersecano il piano *P* ma nei suoi punti ideali (o infinitamente distanti), giacenti nelle direzioni corrispondenti, e che tale piano del fascio interseca il piano *P* lungo una linea ideale (o infinitamente distante). Il piano *P*, completato da questi punti ideali e dalla linea ideale, è chiamato piano completato o proiettivo. Lo denoteremo con P^*. Gli insiemi di linee e piani del

fascio S vengono quindi mappati uno a uno sugli insiemi di punti (reali e ideali) e linee (reali e ideali) di questo piano proiettivo P^*.

Conveniamo quindi nel dire che un punto (reale o ideale) giace su una linea (reale o ideale) del piano proiettivo P^* se la corrispondente retta del fascio giace nel corrispondente piano del fascio. Da questo punto di vista, due linee qualsiasi del piano proiettivo si intersecano (in un punto reale o ideale), poiché due piani qualsiasi del fascio si intersecano lungo una retta del fascio. Da ciò consegue, tra l'altro, che la linea ideale consiste semplicemente nell'insieme di tutti i punti ideali.

In sostanza, il completamento del piano con i suoi elementi ideali significa che usiamo questo piano come una sezione trasversale per studiare il fascio di tutte le linee e i piani che passano per un punto.

Mappature proiettive; il teorema fondamentale. Per mappatura proiettiva intendiamo la mappatura di un piano proiettivo P^* su un altro piano proiettivo $P^{*\prime}$ (che può anche coincidere con il piano P^*, nel qual caso si parla di trasformazione proiettiva del piano P^*), che, prima di tutto, è una corrispondenza biunivoca, e in secondo luogo, è tale che insiemi collineari di punti del piano P^* vanno in insiemi collineari di punti del piano $P^{*\prime}$, e viceversa. (Qui per punti e rette intendiamo sempre punti e rette reali così come ideali.)

È chiaro che due proiezioni prospettiche arbitrarie di uno stesso piano P^* su un piano Π^* possono essere ottenute l'una dall'altra mediante trasformazioni proiettive.

Infatti, 1, i loro punti (reali o ideali) sono in corrispondenza biunivoca con i punti (reali o ideali) del piano proiettivo P^* e di conseguenza tra loro, e 2, punti collineari della prima proiezione corrispondono a punti collineari del piano P^* e di conseguenza anche della seconda proiezione, e viceversa. Pertanto, il suddetto teorema della teoria delle prospettive è una conseguenza diretta del seguente teorema sulle trasformazioni proiettive: se durante una trasformazione proiettiva del piano Π^*, quattro dei suoi punti A, B, C, D, formando una quadrupla di posizione generale rimangono fissi, allora tutti i suoi punti rimangono fissi.

Descriviamo l'idea della dimostrazione di questo teorema mediante la cosiddetta rete di Möbius.

Notiamo che (1) se sotto una trasformazione proiettiva due punti rimangono fissi, allora la linea che li attraversa è mappata su se stessa, e (2) se due linee sono mappate su se stesse, allora il punto della loro intersezione rimane fisso. Pertanto, dal fatto che i punti A, B, C, D del piano Π^* rimangono fissi, ne consegue a loro volta che anche i punti E, F, G, H, K, L, ecc. rimangono fissi (figura 81). La costruzione di tali punti può essere proseguita unendo i punti già ottenuti. Questa è la cosiddetta rete di Möbius. Continuando la sua costruzione, possiamo trovare punti densamente distribuiti a nostro piacimento. Si può dimostrare che l'insieme di questi nodi ricopre ovunque densamente l'intero piano. Pertanto, se assumiamo ulteriormente la continuità di una trasformazione proiettiva (che in realtà consegue già dalla sua definizione, anche se la dimostrazione di questo fatto non è facile), il risultato è che se sotto una trasformazione proiettiva del piano Π^* i punti A, B, C, D rimangono fissi, allora tutti i punti del piano Π^* rimangono fissi.

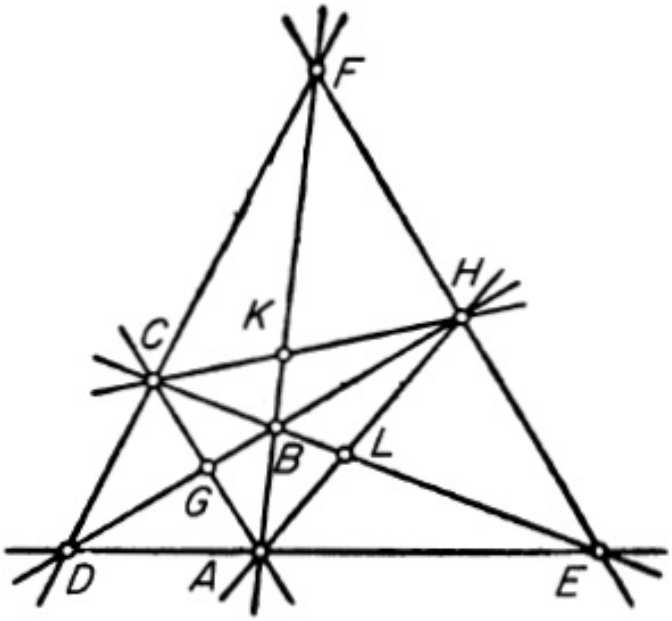


Figura 81

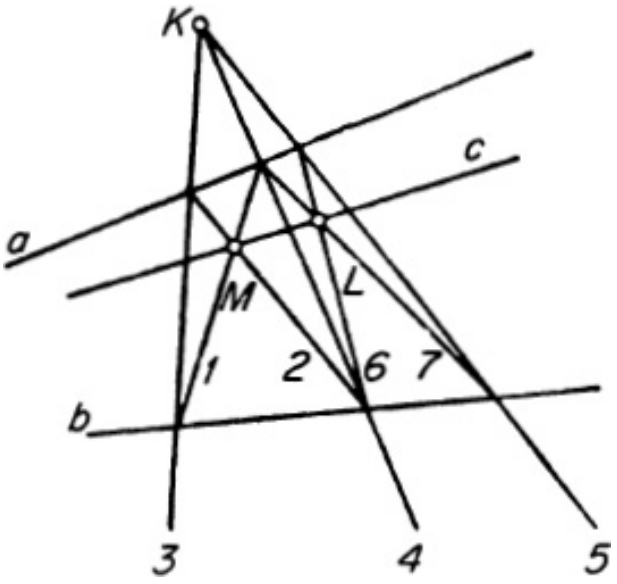


Figura 82.

Geometria proiettiva. Per geometria proiettiva bidimensionale intendiamo l'insieme dei teoremi su quelle proprietà delle figure nel piano proiettivo, cioè il piano ordinario completato da elementi ideali, che non cambiano sotto trasformazioni proiettive arbitrarie.

Ecco un esempio di problema di geometria proiettiva. Date due rette *a* e *b* ed un punto *M* (figura 82), il problema è costruire la retta *c* passante per il punto *M* e per il punto di intersezione delle rette *a* e *b*, non utilizzando tale punto di intersezione (come potrebbe essere necessario, se questo punto è molto distante). Se attraverso il punto *M* tracciamo le due secanti 1 e 2 e poi le linee 3 e 4 attraverso i punti della loro intersezione con le linee *a* e *b*, otteniamo il punto *K*. Tracciamo per essa la linea 5 e le secanti 6 e 7; allora si può dimostrare che la linea *c* passante per il punto *L* di intersezione delle linee 6 e 7 e il punto *M*, è la linea desiderata.

Dalla teoria delle sezioni coniche risulta (figura 83) che l'ellisse, l'iperbole e la parabola sono proiezioni prospettiche l'una dell'altra, e inoltre tutte sono proiezioni prospettiche del cerchio.

Se consideriamo le proiezioni prospettiche come trasformazioni proiettive dei piani proiettivi P^* e $P^{*\prime}$ uno sull'altro, sovrapponendo questi piani otteniamo il risultato che tutte le ellissi, iperboli e parabole sono trasformazioni proiettive del cerchio. La differenza tra loro è che le immagini proiettive del cerchio sotto trasformazioni in cui una linea che non interseca il cerchio è mappata nella linea infinitamente distante sono ellissi; d'altra parte, se una linea tangente al cerchio viene mappata nella linea infinitamente distante, allora si ottiene una parabola, e se una secante, allora un'iperbole (figura 84).

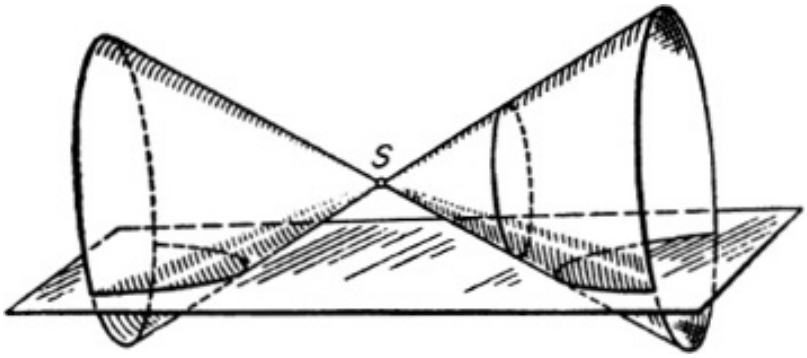


Figura 83.

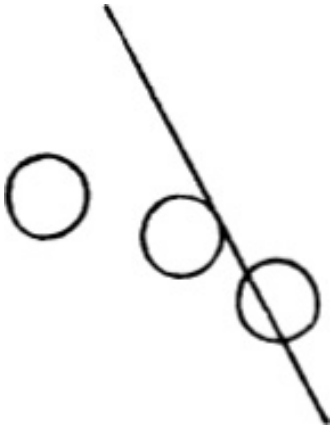


Figura 84.

La notazione delle trasformazioni proiettive nelle formule. Se sul piano P^* prendiamo un sistema di coordinate cartesiane ordinarie, allora, come si può vedere, le formule per le trasformazioni proiettive del piano sono le seguenti

$$x' = \frac{a_1x + b_1y + c_1}{a_3x + b_3y + c_3}, \quad y' = \frac{a_2x + b_2y + c_2}{a_3x + b_3y + c_3},$$

dove il determinante

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0,$$

e viceversa.

Se per qualche punto (x, y) i denominatori sono uguali a zero, ciò significa che la sua immagine (x', y') è un punto ideale (infinitamente distante). L'equazione

$$a_3x + b_3y + c_3 = 0$$

rappresenta la linea che sotto la trasformazione proiettiva data va nella linea ideale (infinitamente distante).

§ 14. Trasformazioni di Lorentz

La derivazione delle formule della trasformazione di Lorentz per il moto lungo una retta e nel piano dalla condizione di costanza della velocità della luce. Alla fine del XIX secolo in fisica fu scoperta una contraddizione fondamentale. Il noto esperimento di Michelson, in cui fu misurata la velocità della luce (che è di circa 300.000 km/sec) nella direzione del moto della Terra lungo la sua orbita attorno al Sole (la velocità della Terra è di circa 30 km/sec) e perpendicolare a questa direzione, dimostrò inconfutabilmente che tutti i corpi in movimento in natura, anche se si muovono nel vuoto, sono contratti nella direzione del moto. La teoria di questa contrazione fu studiata in dettaglio dal fisico olandese Lorentz. Egli dimostrò che questa contrazione è tanto maggiore quanto più la velocità del corpo in movimento si avvicina alla velocità della luce nel vuoto, e ad una velocità pari alla velocità della luce la contrazione diventa infinita. Lorentz derivò le formule per questa contrazione. Ma poco dopo il fisico Einstein introdusse in questo problema un punto di vista completamente diverso, al quale Poincaré era già vicino. Einstein sosteneva quanto segue. Se assumiamo che per la propagazione della luce, come per il moto ordinario di un corpo materiale, valga la legge di composizione delle velocità di Galileo, allora la velocità della luce è $c' = c + v$, dove v è la velocità dell'osservatore che si muove verso la sorgente della luce, e c è la velocità della luce per un osservatore stazionario. Dall'esperimento di Michelson segue che $c' = c$. La legge $c' = c + v$ si basa sulla trasformazione

$$\begin{aligned} x' &= x + v_x t, \\ t' &= t, \end{aligned} \qquad (25)$$

che collega la coordinata x di un punto relativo ad un sistema di coordinate I con la sua coordinata x' relativa ad un sistema di coordinate II che ha i suoi assi paralleli agli assi del sistema I e che si muove parallelo all'asse Ox con velocità v_x rispetto al sistema I. Chiaramente queste sono le formule, come dice Einstein, che devono essere cambiate.

Si può dimostrare, come ha fatto di recente, ad esempio, A. D. Aleksandrov, che dall'uguaglianza della velocità della luce in entrambi i sistemi di coordinate x, y, z, t e x', y', z', t' ne consegue già che le formule di trasformazione dalle coordinate x, y, z, t alle coordinate x', y', z', t' sono lineari e omogenee, cioè hanno la forma

$$\begin{aligned} x' &= a_1x + b_1y + c_1z + d_1t, \\ y' &= a_2x + b_2y + c_2z + d_2t, \\ z' &= a_3x + b_3y + c_3z + d_3t, \\ t' &= a_4x + b_4y + c_4z + d_4t. \end{aligned} \qquad (26)$$

Da altre considerazioni si può dimostrare che il loro determinante[7] è uguale all'unità.

Se un punto nel sistema I si muove rettilineo e uniformemente in una data direzione arbitraria con la velocità della luce c, allora $x = v_x t$, $y = v_y t$, $z = v_z t$ e $v_x^2 + v_y^2 + v_z^2 = c^2$, da cui

$$x^2 + y^2 + z^2 - c^2t^2 = 0. \qquad (27)$$

Ma secondo l'esperimento di Michelson anche questo punto del sistema II si muove necessariamente con la stessa velocità della luce c, per cui è necessario anche che

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = 0.$$

Di conseguenza le formule (26) non sono solo trasformazioni arbitrarie lineari, omogenee e con determinante pari a 1, ma devono allo stesso tempo soddisfare la condizione che se le coordinate x, y, z, t sono tali che

$$x^2 + y^2 + z^2 - c^2t^2 = 0,$$

allora anche le coordinate trasformate x', y', z', t' devono soddisfare questa equazione. Tali trasformazioni (26) sono chiamate trasformazioni di Lorentz.

Consideriamo innanzitutto il caso più semplice, quando il punto si muove lungo l'asse Ox. In questo caso le formule (26) hanno la forma

$$\begin{aligned} x' &= a_1x + d_1t, \\ t' &= a_2x + d_2t, \end{aligned} \qquad (26')$$

ed equazione (27)

$$x^2 - c^2t^2 = 0. \qquad (27')$$

Introduciamo la notazione $ct = u$, quando le formule (26′) e l'equazione (27′) assumono la forma

$$x' = a_1 x + \frac{d_1}{c} u,$$
$$u' = a_2 c x + \frac{d_2 c}{c} u \qquad (26_1)$$

E

$$x^2 - u^2 = 0.$$

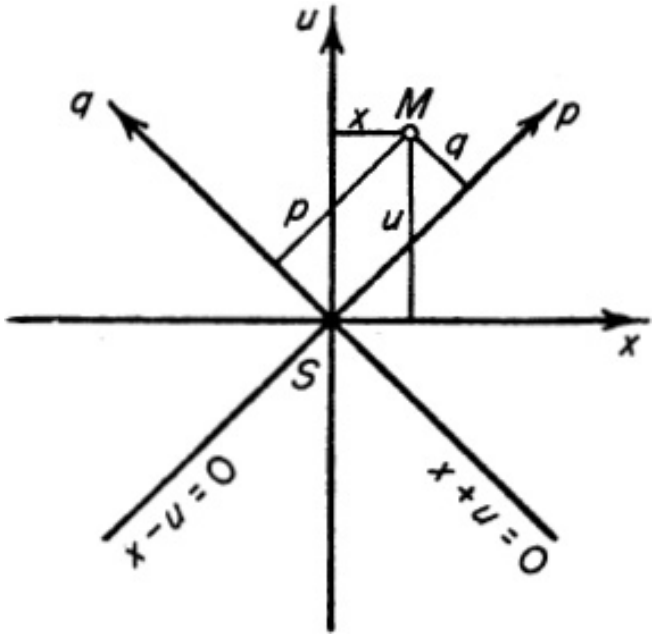


Figura 85.

Troviamo le forme esplicite delle formule (26_1). Consideriamo x e u come un sistema rettangolare cartesiano nel piano, ovvero consideriamo il problema geometricamente; allora possiamo considerare le formule (26_1) come quelle di una trasformazione affine del piano Oxu (il cui determinante, come si è visto, è pari a 1). Indicheremo questa trasformazione con L. Se, come assumiamo, $x^2 - u^2 = 0$ implica $x'^2 - u'^2 = 0$, allora questa trasformazione traduce le rette che si intersecano[E1]

$$x^2 - u^2 = 0$$

in se stesse. La trasformazione L è quindi una combinazione di una contrazione e di un'espansione con uno stesso fattore τ secondo queste linee.

Dalla figura 85 otteniamo

$$x = \frac{p}{\sqrt{2}} - \frac{q}{\sqrt{2}},$$
$$u = \frac{p}{\sqrt{2}} + \frac{q}{\sqrt{2}}.$$

Ma dopo la trasformazione L, i numeri p e q andranno in $p' = p/\tau$ e $q' = q\tau$, quindi

$$x'\sqrt{2} = \frac{p}{\tau} - q\tau,$$

$$u'\sqrt{2} = \frac{p}{\tau} + q\tau.$$

Esprimendo p e q in termini di x e u dalla prima coppia di equazioni, sostituendo nella seconda e semplificando, otteniamo

$$x' = \frac{x - \dfrac{\tau^2 - 1}{\tau^2 + 1}\,ct}{\dfrac{2\tau}{\tau^2 + 1}}, \quad t' = \frac{t - \dfrac{1}{c}\dfrac{\tau^2 - 1}{\tau^2 + 1}\,x}{\dfrac{2\tau}{\tau^2 + 1}},$$

oppure, ponendo $(\tau^2 - 1)/(\tau^2 + 1)\, c = v$, abbiamo

$$x' = \frac{x - vt}{\sqrt{1 - (v/c)^2}}, \quad t = \frac{t - (vx/c^2)}{\sqrt{1 - (v/c)^2}},$$

che sono le famose formule di Lorentz.

In particolare, se prendiamo $x = 0$, cioè se consideriamo il moto dell'origine del sistema di coordinate I, otteniamo

$$x' = \frac{-vt}{\sqrt{1 - (v/c)^2}}, \quad t' = \frac{t}{\sqrt{1 - (v/c)^2}},$$

oppure $x' = -vt'$, da cui ovviamente v è la velocità di movimento del sistema di coordinate II rispetto al sistema I.

Supponiamo, ad esempio, che ci siano dati due punti sull'asse Ox con coordinate x_1 e x_2 rispetto al sistema I, in modo che la distanza tra loro rispetto al sistema I sia $r = |\, x_1 - x_2 \,|$. Vediamo qual è la distanza tra loro per un osservatore collegato al sistema II. Abbiamo

$$x_1' = \frac{x_1 - vt}{\sqrt{1 - (v/c)^2}}, \quad x_2' = \frac{x_2 - vt}{\sqrt{1 - (v/c)^2}},$$

da cui

$$r' = |\, x_1' - x_2' \,| = \frac{|\, x_1 - x_2 \,|}{\sqrt{1 - (v/c)^2}}.$$

Il fattore $\sqrt{1 - (v/c)^2}$ è esattamente il coefficiente della contrazione di Lorentz. Poiché c è molto grande, questo coefficiente è molto vicino a 1 per v moderatamente grande e quindi la contrazione non è significativa. Ma particelle elementari come gli elettroni o i positroni si muovono spesso con velocità paragonabili a quella della luce, e quindi nello studio del loro

movimento è necessario tenere conto di questa contrazione o, come si suol dire, considerare l'effetto relativistico.

Passiamo ora al caso successivo per complessità, ovvero quando il punto si muove nel piano *Oxy*. In questo caso le trasformazioni (26) avranno la forma

$$\begin{aligned} x' &= a_1x + b_1y + d_1t, \\ y' &= a_2x + b_2y + d_2t, \\ t' &= a_3x + b_3y + d_3t, \end{aligned} \qquad (26'')$$

Dove

$$\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} = 1,$$

e l'equazione (27) sarà

$$x^2 + y^2 - c^2t^2 = 0. \qquad (27'')$$

Queste sono le formule di Lorentz per il movimento nel piano *Oxy*.

Ancora una volta mettiamo ct = *u*. Quindi le trasformazioni (26′′) possono essere riscritte come

$$\begin{aligned} x' &= a_1x + b_1y + \frac{d_1}{c}u, \\ y' &= a_2x + b_2y + \frac{d_2}{c}u, \\ u' &= a_3cx + b_3cy + \frac{d_3c}{c}u, \end{aligned} \qquad (26_2)$$

dove il determinante sarà nuovamente uguale a uno, e l'equazione (27′′) assumerà la forma più semplice

$$x^2 + y^2 - u^2 = 0. \qquad (27_2)$$

Considereremo *x*, *y*, *u* come le coordinate cartesiane rettangolari di un punto nello spazio tridimensionale ordinario e considereremo le formule (26_2) come quelle delle trasformazioni affini di questo spazio. L'equazione (27_2) rappresenta un cono circolare retto *K* con un angolo al vertice di 90° (figura 86).

Dal punto di vista di questa interpretazione geometrica (la chiamiamo geometrica perché qui consideriamo *u* = ct semplicemente come una coordinata spaziale) di una trasformazione di Lorentz, l'insieme dei movimenti nel piano è identico all'insieme di tutte le trasformazioni

equiaffini (cioè affini e che preservano il volume) dello spazio che mappa il cono K su se stesso.

Consideriamo alcune speciali trasformazioni di Lorentz.

1. È chiaro che qualsiasi semplice rotazione rigida attorno all'asse del cono K attraverso un angolo ω è una trasformazione equiaffine dello spazio, che mappa il cono K in se stesso, cioè è una trasformazione di Lorentz speciale. La denoteremo con ω.

2. Le riflessioni dello spazio in un piano arbitrario π passante per l'asse del cono K sono chiaramente anche trasformazioni di Lorentz. Le indicheremo con π.

3. Consideriamo infine la seguente trasformazione (figura 87). Siano v e w una qualsiasi coppia di generatrici opposte del cono, e siano P e Q i piani tangenti al cono lungo queste generatrici.

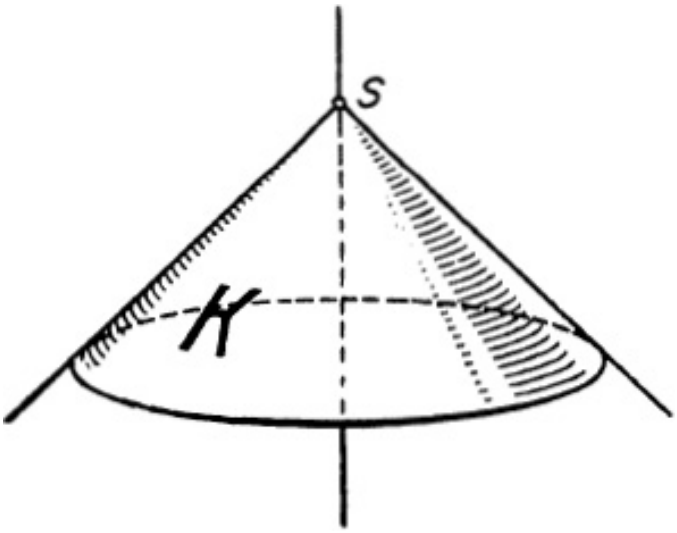


Figura 86.

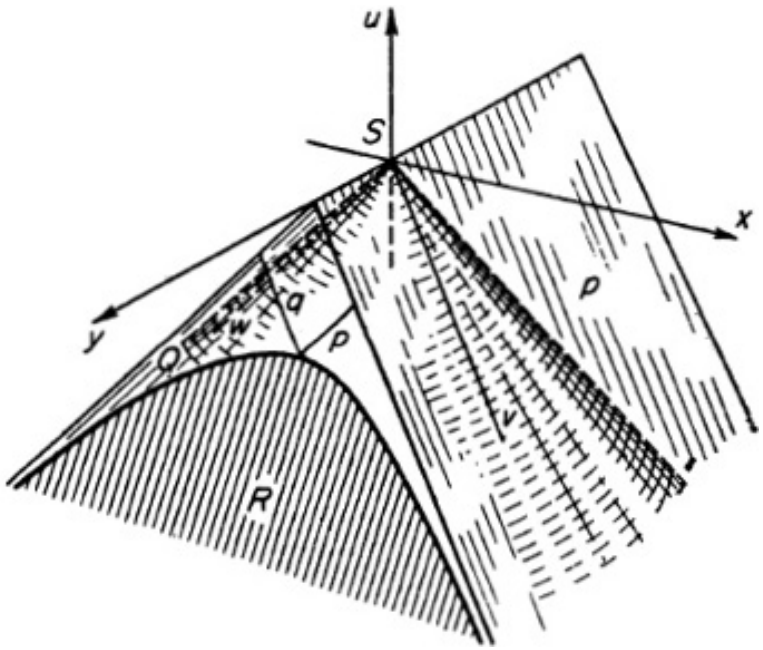


Figura 87.

I piani sono tra loro perpendicolari. Facciamo una contrazione dello spazio rispetto al piano *P* ed una sua espansione con lo stesso coefficiente dal piano *Q*, o viceversa. Ad esempio, contraiamo lo spazio di un fattore tre rispetto al piano *P* e lo espandiamo anch'esso di un fattore tre dal piano *Q*. Tale trasformazione dello spazio è chiaramente anche affine e preserva tutti i volumi. Lo denoteremo con *L*. Mostriamo che questa trasformazione mappa il cono *K* in se stesso. Poiché il cono *K* ha come asse di rivoluzione l'asse *u*, qualsiasi figura può essere ruotata in modo tale che le generatrici *v* e *w* si trovino, ad esempio, nel piano Sxu. Pertanto è sufficiente effettuare la dimostrazione per questo caso.

Per la dimostrazione intersechiamo il cono *K* con un piano arbitrario *R* parallelo al piano Sxu. L'equazione di questo piano è $y = b$, dove b è una costante. Sostituendo questo valore nell'equazione del cono *K*, otteniamo

$$x^2 - u^2 = -b^2.$$

Questa è l'equazione di un'iperbole per la quale le rette di intersezione del piano *R* con i piani *P* e *Q* sono esattamente gli asintoti. Ma poiché per un punto di tale iperbole è caratteristico che il prodotto delle distanze p e q rispetto agli asintoti, cioè ai piani *P* e *Q*, sia costante, durante la trasformazione *L* tutti i punti di questa iperbole rimangono sulla stessa iperbole e l'iperbole è mappata su se stessa. Ma tutta la superficie del cono *K* è costituita da tali iperboli, e quindi durante la trasformazione *L* dello spazio il cono *K* viene rinviato in se stesso. Questa trasformazione *L* è quindi anche una trasformazione di Lorentz.

Poiché nelle trasformazioni affini le rette vanno in rette e le linee intersecanti entrano in linee intersecanti, quindi un fascio *S* di rette sotto qualsiasi trasformazione di Lorentz viene mappato uno a uno su se stesso. Inoltre, sotto trasformazioni affini dello spazio tutti i piani diventano piani, così che sotto queste trasformazioni del fascio *S* su se stesso si ottiene una trasformazione proiettiva del fascio. Se intersechiamo questo fascio con un piano Π

perpendicolare all'asse del cono *K*, che nel suo insieme non viene alterato dalla trasformazione di Lorentz dello spazio data, ed estendiamo questo piano al piano proiettivo Π^* e quindi tracciamo i punti di intersezione delle linee del fascio *S* con il piano Π^*, abbiamo il risultato che le trasformazioni di Lorentz del fascio produrranno contemporaneamente trasformazioni proiettive Λ del piano Π^* e queste ultime trasformeranno in se stesso il cerchio α, in cui il piano Π^* interseca la parte interna del cono *K*. Per analizzare le proprietà delle trasformazioni di Lorentz, è più semplice esaminare queste trasformazioni proiettive Λ del cerchio α in se stesso.

Trasformazioni proiettive di un cerchio in se stesso. Un punto, una semiretta che ne esce e uno dei semipiani tagliati dall'intera linea saranno chiamati “riferimento” del piano Π^* (da non confondere con il riferimento di coordinate, §11). Mostriamo (figura 88) che se prendiamo due riferimenti arbitrari *M* e *M′* contenenti punti interni del cerchio α, allora mediante le trasformazioni *L*, ω, π possiamo inviare uno di questi riferimenti nell'altro. Per questo è sufficiente effettuare le trasformazioni $\Lambda = L_1 \cdot \omega \cdot L_2^{-1}$ (oppure $\Lambda = L_1 \cdot \omega \cdot \pi \cdot L_2^{-1}$). La trasformazione L_1 invia il primo riferimento *M* al centro *O* del cerchio α, la trasformazione ω lo ruota quanto necessario, ed infine la trasformazione L_2^{-1} lo porta in coincidenza con il secondo riferimento *M′*.

Mostriamo, inoltre, che esiste una sola trasformazione Λ che traduce un dato riferimento *M* in un dato riferimento *M′*. Per fare ciò osserviamo innanzitutto che se ci fossero due trasformazioni Λ_1 e Λ_2 che inviano il riferimento *M* nel riferimento *M′*, allora la trasformazione $\Lambda = \Lambda_1\Lambda_2^{-1}$ non sarebbe la trasformazione di identità Λ e invierebbe il riferimento *M* in se stesso. Pertanto è sufficiente mostrare che se una trasformazione Λ manda in sé il riferimento *M*, allora è identità, cioè lascia fissi tutti i punti del piano del cerchio α.

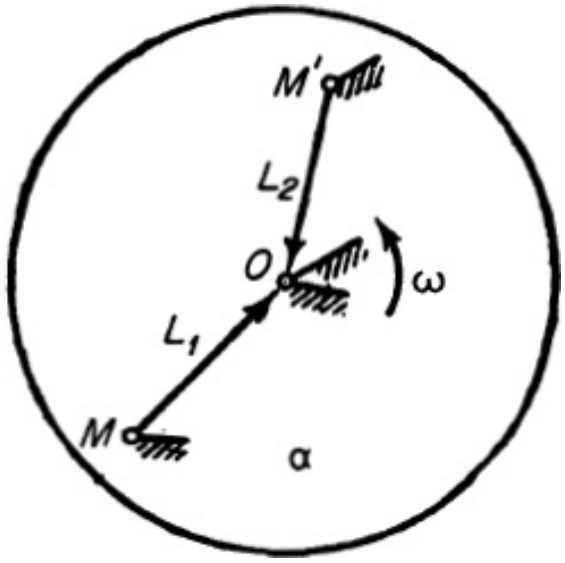


Figura 88.

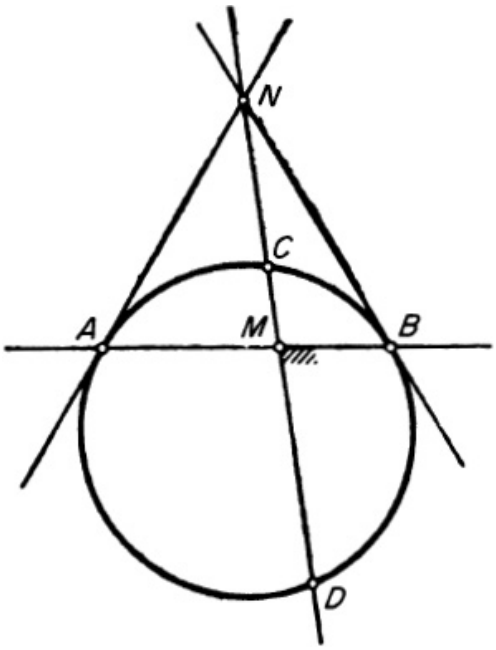


Figura 89.

Mostriamo questo. Supponiamo che la trasformazione Λ invii in sé il riferimento M (figura 89). Poi mappa in sé la linea AB di questo riferimento, ma poiché manda in sé la circonferenza del cerchio α, lascia perciò fissi i punti A e B, oppure li scambia. Quest'ultima cosa però è impossibile, poiché la semiretta del riferimento è mappata su se stessa. Disegniamo le tangenti nei punti A e B al cerchio α. Sono mappate su se stesse, poiché se tale tangente fosse mappata in una secante $A\bar{A}$, la trasformazione inversa invierebbe i diversi punti A e $\bar{A}$ del cerchio α nell'unico punto A. Ma le Λ sono trasformazioni proiettive, e di conseguenza biunivoche. Poiché nelle trasformazioni Λ queste tangenti vanno in se stesse, quindi il punto N della loro intersezione rimane fisso, e di conseguenza la linea MN viene mappata in se stessa. Dal fatto che il semipiano del riferimento M è mappato su se stesso, concludiamo come sopra che i punti C e D non sono scambiati, ma rimangono fissi. Quindi, sotto la trasformazione proiettiva data Λ del piano proiettivo Π^*, quattro dei suoi punti A, B, C, D, di cui non ce ne sono tre che si trovano sulla stessa linea, rimangono fissi. Secondo il teorema di unicità delle trasformazioni proiettive questa è la trasformazione dell'identità.

Successivamente, nel §5 del capitolo XVII si dimostrerà che utilizzando queste proprietà del gruppo di Lorentz è facile costruire un modello della geometria piana di Lobačevskij, e se consideriamo le trasformazioni di Lorentz per il caso generale del moto di un punto nello spazio, allora possiamo fare lo stesso per la geometria spaziale di Lobačevskij, dimostrandone così la non contraddittorietà.

Vediamo che la teoria delle trasformazioni di Lorentz, la geometria proiettiva e la teoria della prospettiva e la geometria non euclidea sono strettamente correlate tra loro. Si scopre che esiste anche un'altra teoria strettamente correlata ad essi, vale a dire le cosiddette trasformazioni conformi nella teoria delle funzioni di variabile complessa, che risolvono importanti problemi di fisica matematica come la distribuzione della temperatura in una

piastra riscaldata, il flusso d'aria attorno all'ala di un aereo, la distribuzione della carica in un campo elettrostatico piano, i problemi dell'elasticità nel piano e molti altri.

Conclusione

La geometria analitica è un metodo assolutamente indispensabile per l'indagine di altri rami della matematica, della fisica e di altre scienze naturali. Pertanto viene studiata non solo nelle università ma in tutti gli istituti superiori di formazione tecnica, e anche in alcune scuole professionali. Si tratta anche di stabilire se non sia il caso di includere nei corsi delle scuole superiori una trattazione abbastanza dettagliata degli elementi di geometria analitica.

Coordinate varie. Gli elementi essenziali del concetto di geometria analitica, come abbiamo visto, sono il metodo delle coordinate e lo studio delle equazioni che collegano queste coordinate. Oltre alle coordinate cartesiane se ne possono considerare altre diverse. Ad esempio, nel piano, possiamo scegliere un punto P (il cosiddetto polo) e un raggio che da esso proviene (l'asse polare) e determinare la posizione di un punto M mediante la lunghezza ρ del raggio polare dal polo al punto e il valore ω dell'angolo formato da questo raggio con l'asse polare (figura 90).

In particolare l'ellisse, l'iperbole o la parabola, se prendiamo per polo

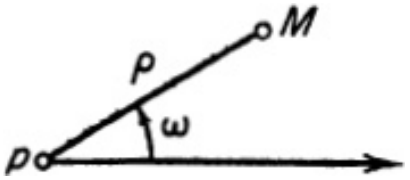


Figura 90.

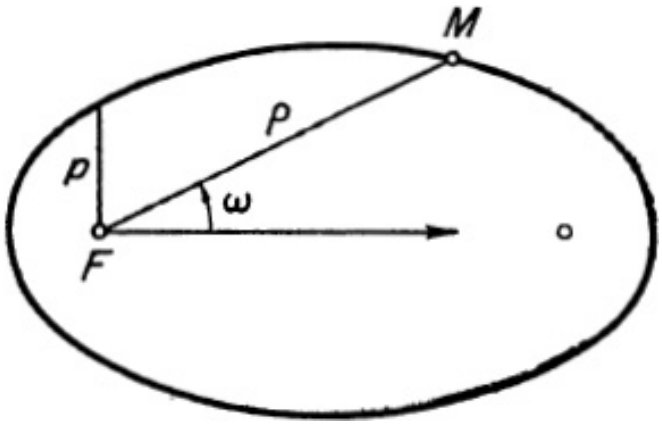


Figura 91.

un fuoco, e per l'asse polare il raggio che passa dal fuoco lungo l'asse di simmetria alla parte opposta al vertice più vicino (figura 91), hanno la stessa equazione

$$\rho = \frac{p}{1 - \epsilon \cos \omega},$$

dove ϵ è l'eccentricità della curva e p è il suo cosiddetto parametro. Questa equazione è di grande importanza in astronomia. Infatti con il suo aiuto si derivò il risultato, dalla legge di inerzia e dalla legge di gravitazione universale, che i pianeti ruotano attorno al Sole secondo ellissi.

Sono ben note le coordinate geografiche, latitudine e longitudine, con le quali viene data la posizione di un punto su una sfera.

Analogamente, possiamo prendere una rete di coordinate su una superficie arbitraria, come si fa in geometria differenziale (vedi Capitolo VII), ecc.

Geometria analitica a più dimensioni e a infinite dimensioni; geometria algebrica. Sembrerebbe che nel XIX secolo la geometria analitica abbia avuto uno sviluppo così immenso, descritto prima in modo generale, e abbia prodotto così tante idee, che si sarebbe necessariamente esaurita, ma non è così. In tempi molto recenti si sono sviluppate rapidamente due nuove ed estese branche della matematica che hanno esteso i concetti della geometria analitica, vale a dire la cosiddetta analisi funzionale e la geometria algebrica generale. È vero che entrambe rappresentano solo a metà una diretta continuazione della geometria analitica classica: gran parte dell'analisi funzionale è analisi, e nella geometria algebrica hanno ampio spazio della teoria delle funzioni e della topologia.

Chiariamo che cosa intendiamo. Già a metà del secolo scorso i matematici avevano cominciato a considerare la geometria analitica a quattro dimensioni e, più in generale, a n dimensioni, studiando quei problemi algebrici che generalizzano direttamente al caso di quattro o n incognite i problemi della geometria analitica a due e a tre dimensioni. Alla fine del XIX secolo alcuni eminenti analisti giunsero all'idea che, per gli scopi dell'analisi e della fisica matematica, fosse importante considerare una geometria analitica a infinite dimensioni.

A prima vista può sembrare che gli spazi n-dimensionali o addirittura quadridimensionali sembrino inverosimili finzioni matematiche, quindi lo stesso si può dire anche di uno spazio infinitamente dimensionale. Ma non è proprio così. Le argomentazioni riguardanti uno spazio a dimensione infinita non sono affatto difficili. Costituiscono ora un ramo importante della matematica, l'analisi funzionale (vedi Capitolo XIX).

È curioso che la geometria analitica a dimensione infinita abbia applicazioni pratiche importantissime e svolga un ruolo fondamentale nella fisica contemporanea.

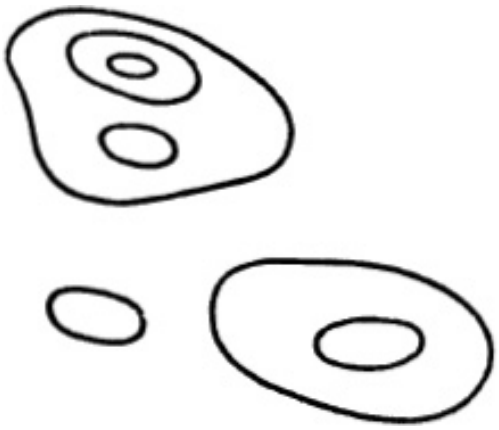


Figura 92.

Quanto alla geometria algebrica, essa è una continuazione più immediata della geometria analitica ordinaria, che a sua volta è solo una parte della geometria algebrica. La geometria algebrica può essere considerata come quella parte della matematica che si occupa di curve, superfici e ipersuperfici, rappresentate in coordinate cartesiane da equazioni algebriche non solo di primo e secondo grado, ma anche di grado superiore. Risulta che in queste indagini è vantaggioso considerare non solo le coordinate reali ma anche quelle complesse, cioè considerare tutto in un cosiddetto spazio complesso. I risultati più importanti in questo campo furono ottenuti nel secolo scorso da Riemann. Come brillante esempio di teoremi sulle curve di ordine superiore segnaliamo un risultato straordinariamente generale di I. G. Petrovskij sul numero di ovali in cui può essere scomposta una curva di ordine n. Petrovskij ha dimostrato che se p è il numero di quegli ovali che non si trovano affatto in altri ovali, o che si trovano in un numero pari di ovali, e m è il numero di quegli ovali che si trovano in un numero dispari di ovali, e se consideriamo solo le curve i cui ovali componenti non si intersecano né con se stessi né fra loro (figura 92), allora

$$p - m \leqslant \frac{3n^2 - 6n}{8} + 1,$$

dove n è l'ordine della curva, cioè il grado dell'equazione con cui è rappresentata la curva.

Questo risultato è tanto più importante in quanto fino ad allora non si sapeva quasi nulla sulla forma generale di una curva di ordine superiore. È senza dubbio uno dei più importanti teoremi generali recenti della geometria analitica.

Letture consigliate

A. A. Albert, Solid analytic geometry, McGraw-Hill, New York, 1949.

J. L. Coolidge, A history of the conic sections and quadratic surfaces, Oxford University Press, Oxford, 1945.

H. S. M. Coxeter, The real projective plane, 2ª ed., Cambridge University Press, New York, 1960.

L. P. Eisenhart, Coordinate geometry, Dover, New York, 1960.

A. Jaeger, Introduction to analytic geometry and linear algebra, Holt, Rinehart and Winston, New York, 1960.

D. J. Struik, Lectures on analytic and projective geometry, Addison-Wesley, Cambridge, Mass., 1953.

Note al capitolo III

[1] A volte, l'equazione non è soddisfatta da nessun punto (x, y) con coordinate reali; altre volte da uno o più di questi punti. In questo caso si dice che la curva è immaginaria o si riduce a punti (vedi §7).

[2] Le eccentricità delle orbite planetarie non sono molto grandi, quindi le orbite dei pianeti sono quasi cerchi.

[3] I coefficienti di xy, x, y saranno indicati non con B, D, E ma con $2B$, $2D$, $2E$ per semplicità delle formule successive.

[4] Come si può facilmente dimostrare.

[5] Sui determinanti vedere il Capitolo XVI.

[6] Invarians in latino significa "immutato".

[7] Vedi capitolo XVI.

Nota editoriale

[E1] Sono corretti i seguenti refusi matematici del testo di partenza: nel §7 la direttrice della parabola $y^2 = 2px$ è $x = -p/2$; nel §10, per l'iperboloide di rivoluzione, il denominatore di y^2 è a^2; nella tabella 1, alla riga delle rette parallele, il primo criterio è $I_2 = 0$; nel §12 il parametro p è quello della parabola; nel §14 l'equazione trasformata del cono di luce contiene $x'^2 - u'^2$, con il segno meno. La figura 50 è recuperata direttamente dal DjVu, poiché nell'edizione elettronica era sostituita da un'altra figura. Le notazioni e la formulazione storica del resto del capitolo sono conservate.

CAPITOLO IV

Algebra: teoria delle equazioni algebriche

B. N. Delone

§ 1. Introduzione

I tratti caratteristici dell'algebra sono ben noti a tutti, poiché le informazioni elementari ma fondamentali che la riguardano vengono già impartite alle scuole superiori. L'algebra si caratterizza innanzitutto per il suo metodo, che prevede l'uso di lettere ed espressioni in lettere, sulle quali si eseguono operazioni secondo leggi definite. Nell'algebra elementare le lettere denotano numeri ordinari, così che le leggi delle operazioni sulle espressioni in lettere si basano sulle leggi generali delle operazioni sui numeri. Ad esempio, la somma non dipende dall'ordine degli addendi, fatto che in algebra si scrive così: $a + b = b + a$; nel moltiplicare la somma di due numeri, possiamo moltiplicare ciascuno dei numeri singolarmente e poi sommare i prodotti così ottenuti: $(a + b)\, c = ac + bc$, ecc.

Se tracciamo la dimostrazione di un teorema algebrico, è facile vedere che essa dipende solo da queste leggi per le operazioni sui numeri e per niente da ciò che rappresentano le lettere.

Il metodo algebrico, cioè il metodo dei calcoli con le lettere, penetra tutta la matematica. Infatti, una parte sostanziale della soluzione di un problema matematico spesso non è altro che un calcolo algebrico più o meno complicato. Inoltre in matematica utilizziamo vari calcoli simbolici in cui le lettere non indicano più numeri ma altre entità, dove le leggi per le operazioni su queste entità possono essere diverse dalle leggi dell'algebra elementare. Ad esempio, in geometria, meccanica e fisica facciamo uso dei vettori e, come è noto, le leggi per le operazioni sui vettori sono in parte le stesse che per i numeri e in parte essenzialmente diverse.

L'importanza del metodo algebrico nella matematica moderna e la gamma delle sue applicazioni sono aumentate notevolmente negli ultimi decenni.

Innanzitutto, le crescenti esigenze della tecnologia ci costringono a ridurre a risultati numerici le soluzioni di difficili problemi di analisi matematica, e ciò di solito risulta fattibile solo dopo l'algebrizzazione di questi problemi, processo che a sua volta crea nuovi e talvolta difficili problemi nell'algebra stessa.

In secondo luogo, alcuni problemi di analisi divennero chiari e comprensibili solo dopo essere stati affrontati con metodi algebrici basati su una profonda generalizzazione (al caso di infinite incognite) della teoria dei sistemi di equazioni di primo grado.

Infine, le parti più avanzate dell'algebra hanno trovato applicazione nella fisica contemporanea: infatti, i concetti fondamentali della meccanica quantistica sono espressi in termini di entità algebriche complicate e non elementari.

Le caratteristiche fondamentali della storia dell'algebra sono le seguenti.

Innanzitutto dobbiamo sottolineare che le nostre idee su cosa sia l'algebra e in cosa consista il suo problema fondamentale sono cambiate due volte: una nella prima metà del secolo scorso, e la seconda volta all'inizio del nostro secolo. Pertanto, l'algebra ha significato in tempi diversi tre cose del tutto diverse. Sotto questo aspetto la storia dell'algebra differisce dalla storia dei tre famosi rami della matematica: geometria analitica, calcolo differenziale e calcolo integrale, che furono forgiati per mano dei loro creatori, Fermat, Cartesio, Newton, Leibniz e altri e in seguito furono rapidamente sviluppati e amplificati, a volte con l'aggiunta di grandi nuove sezioni, ma furono relativamente poco cambiati nel loro carattere fondamentale.

Nell'antichità ogni legge che veniva scoperta per la soluzione di una classe di problemi matematici veniva scritta semplicemente a parole, poiché i calcoli simbolici non erano ancora stati inventati. La parola stessa "algebra" è stata creata dal nome dell'importante lavoro dello scienziato del IX secolo originario della Corasmia, Mohammed Al-Kharizmi (vedi Capitolo I), nelle cui opere fu dedotta la prima legge generale per la soluzione delle equazioni di primo e secondo grado. Tuttavia, l'introduzione della stessa notazione simbolica è solitamente associata al nome di Viète, che iniziò a denotare con lettere non solo le incognite ma anche le quantità date. Anche Cartesio fece molto per lo sviluppo della notazione simbolica, e anche lui, ovviamente, interpretò le lettere come numeri ordinari. È in questo momento che comincia realmente l'algebra come scienza dei calcoli simbolici, delle trasformazioni di formule composte di lettere, delle equazioni algebriche, ecc., in contrasto con l'aritmetica, che opera sempre su numeri concreti. Solo ora i concetti matematici complicati sono diventati chiari e accessibili all'indagine, poiché dando uno sguardo a una formula in lettere, nella maggior parte dei casi è possibile per noi vedere la sua disposizione generale o legge di formazione e sottoporla ad opportune trasformazioni. A quel tempo tutto ciò che in matematica non era né geometria né analisi infinitesimale veniva chiamato algebra. Questo è il primo punto di vista, per così dire di Viète, riguardo all'algebra. Ciò è stato espresso molto chiaramente nel famoso libro "Introduzione all'algebra" di un membro dell'Accademia russa delle scienze, il famoso L. Eulero, scritto nel 1760, cioè 200 anni fa.

Eulero definì l'algebra come la teoria dei calcoli con quantità di vario tipo. La prima parte del suo libro contiene la teoria del calcolo con numeri interi, le frazioni ordinarie, le radici quadrate e cubiche, la teoria dei logaritmi, le progressioni, la teoria dei calcoli con polinomi, la teoria delle serie binomiali di Newton e le sue applicazioni. La seconda parte comprende la teoria delle equazioni di primo grado e dei sistemi di tali equazioni, la teoria delle equazioni quadratiche e delle soluzioni delle equazioni di terzo e quarto grado mediante radicali, nonché un'ampia sezione sui metodi di soluzione di varie equazioni indeterminate in numeri interi. Ad esempio, è stato dimostrato che l'equazione di Fermat $x^3 + y^3 = z^3$ non può essere risolta negli interi non nulli x, y, z.

Alla fine del XVIII e all'inizio del XIX secolo, uno dei problemi dell'algebra cominciò gradualmente ad occupare il posto centrale, vale a dire la teoria della soluzione delle equazioni algebriche, in cui la difficoltà fondamentale è la soluzione di un'equazione algebrica di grado n con un'incognita

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0.$$

Ciò avvenne come naturale conseguenza dell'importanza del problema per tutta la matematica pura e applicata, e anche per la difficoltà e la profondità della maggior parte dei teoremi ad esso connessi.

La formula generale per la soluzione di un'equazione quadratica,

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q},$$

era nota a tutti. Gli algebristi italiani del XVI secolo trovarono regole generali analoghe, sebbene più complicate, per la soluzione di equazioni arbitrarie di terzo e quarto grado. Ulteriori indagini in questa direzione per le equazioni di grado superiore, tuttavia, incontrarono difficoltà insormontabili. I più grandi matematici del XVI, XVII, XVIII e dell'inizio del XIX secolo (Tartaglia, Cardano, Cartesio, Newton, d'Alembert, Tschirnhausen, Bézout, Lagrange, Gauss, Abel, Galois, Lobačevskij, Sturm e altri) crearono un impressionante edificio di teoremi e metodi legati a questo problema. L'algebra in due volumi di Serret (un'opera epocale per i suoi tempi, poiché presentò per la prima volta il culmine della teoria delle equazioni algebriche, cioè la teoria di Galois), apparve a metà del XIX secolo, esattamente 100 anni dopo il testo di Eulero; in esso l'algebra era già definita come la teoria delle equazioni algebriche. Questo è il secondo punto di vista su cosa sia l'algebra.

Nella seconda metà del secolo scorso si è verificato, sulla base delle idee di Galois sulla teoria delle equazioni algebriche, un profondo sviluppo della teoria dei gruppi[1] e della teoria dei numeri algebrici (nella cui creazione gran parte ha avuto il matematico russo E. I. Zolotarev).

Anche in questo secondo periodo, in connessione con gli stessi problemi di soluzione di un'equazione algebrica, e con la teoria delle varietà algebriche di ordine superiore (che allora venivano studiate nella geometria analitica) l'apparato algebrico fu sviluppato in diverse direzioni, ad esempio la teoria dei determinanti e delle matrici, la teoria algebrica delle forme quadratiche e delle trasformazioni lineari, e, in particolare, la teoria degli invarianti. Per quasi tutta la seconda metà dell'Ottocento la teoria degli invarianti fu un tema centrale dell'algebra. A sua volta, lo sviluppo della teoria dei gruppi e della teoria degli invarianti esercitò in questo periodo una grande influenza sullo sviluppo della geometria.[2]

Un nuovo, terzo punto di vista su cosa sia l'algebra è nato principalmente nel seguente collegamento. Nella seconda metà del secolo scorso, nella meccanica, nella fisica e nella stessa matematica, gli scienziati iniziarono sempre più spesso a indagare oggetti per i quali era naturale considerare operazioni di addizione e sottrazione, e talvolta moltiplicazione e

divisione, ma per i quali queste operazioni erano soggette a leggi del tutto diverse da quelle dei numeri razionali.

Abbiamo già parlato dei vettori. Altri tipi di oggetti matematici con leggi delle operazioni diverse possono essere qui menzionati solo: ad esempio matrici, tensori, spinori, numeri ipercomplessi. Tutte queste quantità sono indicate con lettere, ma le loro leggi delle operazioni differiscono l'una dall'altra. Se per un certo insieme di oggetti (indicati con lettere) sono definite determinate operazioni insieme alle leggi o regole che devono soddisfare, allora diciamo che è definito un sistema algebrico. Il terzo punto di vista su cosa sia l'algebra consiste nel considerare l'algebra nel suo insieme come lo studio di vari sistemi algebrici. Questa è la cosiddetta algebra assiomatica o astratta. È astratta perché in un dato passo del calcolo non ci interessa affatto a ciò che denotano le lettere del sistema algebrico, l'unica cosa importante sono gli assiomi o le leggi soddisfatte dalle operazioni; ed è detta assiomatica, perché è costruita esclusivamente a partire dagli assiomi enunciati all'inizio. È come se fossimo tornati, ma a un livello più alto, al primo punto di vista dell'algebra o di Viète, secondo cui l'algebra è la teoria dei calcoli simbolici. Sebbene non faccia alcuna differenza ciò che le lettere denotano e siano importanti solo le regole delle operazioni, è pur sempre vero, naturalmente, che sono interessanti solo quei sistemi algebrici che hanno un grande significato sia nella matematica stessa che nelle sue applicazioni.

La grande quantità di materiale algebrico raccolto nel periodo precedente servì come base effettiva per la costruzione dell'algebra astratta contemporanea.

I primi anni ’30 videro la pubblicazione del famoso libro di van der Waerden “Modern Algebra”, che ha svolto un ruolo importante nella propagazione di questo terzo punto di vista su cosa sia l’algebra. Il testo di A. G. Kuroš sull’algebra è orientato nella stessa direzione.

Nel nostro secolo l'algebra ha trovato profonde applicazioni alla geometria (topologia e teoria dei gruppi di Lie) e, come accennato in precedenza, alla fisica contemporanea, in particolare all'analisi funzionale e alla meccanica quantistica.

Particolarmente importanti al giorno d'oggi sono i problemi della meccanizzazione dei calcoli algebrici mediante varie macchine di calcolo matematico, in particolare macchine elettroniche ad alta velocità. Le questioni legate a questo tipo di matematica computazionale sollevano nuovi problemi distintivi in algebra.

Nel presente lavoro ci sono due capitoli (senza contare il presente) dedicati all'algebra: l'algebra lineare (Capitolo XVI) e la teoria dei gruppi e altri sistemi algebrici (Capitolo XX).

§ 2. Risoluzione algebrica di un’equazione

Un'equazione algebrica del grado *n* con un'incognita è un'equazione della forma

$$x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n = 0,$$

dove a_1, a_2, ⋯, a_n, sono coefficienti dati.[3]

Equazioni di primo e secondo grado.

Se l'equazione è di primo grado, allora ha la forma

$$x + a = 0$$

e si risolve immediatamente

$$x = -a.$$

L'equazione di secondo grado

$$x^2 + px + q = 0$$

era già risolta nell’antichità. La sua soluzione è molto semplice: se trasferiamo q con il segno opposto a destra e poi aggiungiamo $p^2/4$ su entrambi i lati abbiamo

$$x^2 + px + \frac{p^2}{4} = \frac{p^2}{4} - q.$$

Ma

$$x^2 + px + \frac{p^2}{4} = \left(x + \frac{p}{2}\right)^2,$$

quindi

$$x + \frac{p}{2} = \pm\sqrt{\frac{p^2}{4} - q},$$

da cui si ottiene la nota formula per le soluzioni di un'equazione quadratica

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}.$$

Equazione di terzo grado.

La situazione era completamente diversa per le equazioni di grado superiore a 2. Già l'equazione generale di terzo grado richiedeva considerazioni abbastanza profonde e resisteva a tutti gli sforzi dei matematici dell'antichità. Fu risolta solo all'inizio del 1500, nell'era del Rinascimento in Italia, dal matematico italiano Scipione del Ferro. Del Ferro, secondo l'uso del suo tempo, non pubblicò le proprie scoperte ma le comunicò ad un suo allievo. Dopo la morte di del Ferro questo allievo sfidò in gara uno dei grandi matematici italiani Tartaglia e gli propose per la soluzione una serie di equazioni di terzo grado. Tartaglia (1500-1557) accettò la sfida e otto giorni prima della fine della competizione trovò un metodo per risolvere qualsiasi equazione cubica della forma $x^3 + px + q = 0$.

In due ore risolse tutti i problemi del suo avversario. Un professore di fisica e matematica a Milano, Cardano (1501-1576), venendo a conoscenza delle scoperte di Tartaglia, iniziò a supplicare Tartaglia di informarlo del suo segreto. Tartaglia alla fine acconsentì, ma a condizione che Cardano mantenesse il suo metodo profondamente segreto. Cardano venne meno alla sua promessa e pubblicò il risultato di Tartaglia nella sua opera "La grande arte" (Ars Magna).

La formula per la soluzione di un'equazione cubica da allora venne chiamata formula di Cardano, anche se sarebbe corretto chiamarla formula di Tartaglia.

La formula di Cardano si ricava come segue.

In primo luogo, la soluzione dell'equazione cubica generale

$$y^3 + ay^2 + by + c = 0 \qquad (1)$$

può essere facilmente ridotta alla soluzione dell'equazione cubica della forma

$$x^3 + px + q = 0, \qquad (2)$$

non contenente un termine con il quadrato dell'incognita. Per fare ciò è sufficiente porre $y = x - a/3$. Infatti, sostituendo questa espressione nell'equazione (1) e rimuovendo le parentesi, otteniamo

$$\left(x - \frac{a}{3}\right)^3 + a\left(x - \frac{a}{3}\right)^2 + b\left(x - \frac{a}{3}\right) + c = x^3 - 3x^2\frac{a}{3} + \cdots + ax^2 + \cdots,$$

dove i punti indicano quei termini in cui x viene elevato alla prima potenza o non compare affatto. Vediamo che i termini contenenti x^2 si annullano a vicenda.

Consideriamo ora la seguente equazione

$$x^3 + px + q = 0.$$

Poniamo $x = u + v$, cioè al posto di un'incognita ne mettiamo due, u e v, trasformando così l'intero problema in un problema con due incognite.

Abbiamo

$$(u + v)^3 + p(u + v) + q = 0,$$

O

$$u^3 + v^3 + q + (3uv + p)(u + v) = 0.$$

Qualunque sia la somma dei due numeri $u + v$, è sempre possibile richiedere che il loro prodotto uv sia uguale ad una certa quantità data in anticipo. Se $u + v = A$, e imponiamo uv = B, allora poiché $v = A - u$, otteniamo

$$u(A - u) = B,$$

per cui è sufficiente che u sia una soluzione dell'equazione quadratica

$$u^2 - Au + B = 0,$$

e sappiamo che ogni equazione quadratica ha radici reali o complesse, date dalla nota formula. Nel nostro caso, $u + v$ è uguale alla radice desiderata x della nostra equazione cubica e imponiamo che

$$uv = -\frac{p}{3},$$

cioè che 3uv + p = 0. Con questa scelta di u e v otteniamo

$$\begin{aligned} u^3 + v^3 + q &= 0, \\ 3uv + p &= 0. \end{aligned} \qquad (3)$$

Di conseguenza, se troviamo i numeri u e v, che soddisfano questo sistema di equazioni, allora il numero $x = u + v$ sarà la radice della nostra equazione.

Dal sistema (3) è facile formare un'equazione quadratica le cui radici saranno u^3 e v^3. Infatti, esso dà

$$\begin{aligned} u^3 + v^3 &= -q, \\ u^3v^3 &= -\frac{p^3}{27}, \end{aligned}$$

e, di conseguenza per un teorema già utilizzato in precedenza u^3 e v^3 sono le radici dell'equazione quadratica

$$z^2 + qz - \frac{p^3}{27} = 0.$$

Risolvendola con la solita formula, otteniamo

$$u^3 = -\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}, \quad v^3 = -\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}},$$

e, di conseguenza,

$$x = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

questa è la formula di Cardano.

Equazione di quarto grado.

Subito dopo la soluzione dell'equazione cubica, Ferrari (1522-1565) risolse l'equazione generale del quarto grado. Per la soluzione dell'equazione di terzo grado abbiamo visto che la soluzione preliminare dell'equazione quadratica ausiliaria,

$$z^2 + qz - \frac{p^3}{27} = 0,$$

era necessaria, dove $z = u^3$ oppure v^3; analogamente, la soluzione di un'equazione di quarto grado può basarsi sulla soluzione preliminare di un'equazione cubica ausiliaria.

Il metodo di Ferrari consiste in quanto segue. Sia data l'equazione generale di quarto grado

$$x^4 + ax^3 + bx^2 + cx + d = 0.$$

Riscriviamola come:

$$x^4 + ax^3 = -bx^2 - cx - d$$

e aggiungiamo su entrambi i lati $a^2x^2/4$; poi a sinistra otteniamo un quadrato perfetto

$$\left(x^2 + \frac{ax}{2}\right)^2 = \left(\frac{a^2}{4} - b\right)x^2 - cx - d.$$

Aggiungendo ora i termini ad entrambi i lati dell'equazione

$$\left(x^2 + \frac{ax}{2}\right)y + \frac{y^2}{4},$$

dove y è una nuova variabile, sulla quale imponiamo poi una condizione necessaria, a sinistra otteniamo un quadrato perfetto

$$\left(x^2 + \frac{ax}{2} + \frac{y}{2}\right)^2 = \left(\frac{a^2}{4} - b + y\right)x^2 + \left(\frac{ay}{2} - c\right)x + \left(\frac{y^2}{4} - d\right). \qquad (4)$$

Abbiamo quindi ridotto il problema a due incognite.

A destra dell'equazione (4) abbiamo un trinomio quadratico in x, i cui coefficienti dipendono da y. Selezioniamo y tale che questo trinomio sia il quadrato del binomio di primo grado $\alpha x + \beta$.

Affinché il trinomio quadratico $Ax^2 + Bx + C$ sia il quadrato del binomio $\alpha x + \beta$ è sufficiente che

$$B^2 - 4AC = 0.$$

Infatti, se $B^2 - 4AC = 0$, allora

$$Ax^2 + Bx + C = (\sqrt{A}x + \sqrt{C})^2,$$

cioè,

$$Ax^2 + Bx + C = (\alpha x + \beta)^2,$$

Dove

$$\alpha = \sqrt{A}, \quad \beta = \sqrt{C}.$$

Di conseguenza, se scegliamo y in modo che

$$\left(\frac{ay}{2} - c\right)^2 - 4\left(\frac{a^2}{4} - b + y\right)\left(\frac{y^2}{4} - d\right) = 0,$$

allora il membro destro dell'equazione (4) sarà il quadrato perfetto $(\alpha x + \beta)^2$. Togliendo le parentesi, otteniamo un'equazione cubica in y

$$y^3 - by^2 + (ac - 4d)y - [d(a^2 - 4b) + c^2] = 0.$$

Risolvendo questa equazione cubica ausiliaria (ad esempio, con la formula di Cardano) troviamo α e β in termini della sua soluzione y_0, vale a dire

$$\left(x^2 + \frac{ax}{2} + \frac{y_0}{2}\right)^2 = (\alpha x + \beta)^2,$$

da cui

$$x^2 + \frac{ax}{2} + \frac{y_0}{2} = \alpha x + \beta \quad \text{or} \quad x^2 + \frac{ax}{2} + \frac{y_0}{2} = -\alpha x - \beta.$$

Da queste due equazioni quadratiche possiamo trovare tutte e quattro le radici dell'equazione di quarto grado data. Ecco come venivano risolte le equazioni algebriche di terzo e quarto grado dai matematici italiani nel Cinquecento.

Il successo dei matematici italiani produsse un grandissimo effetto. Fu il primo caso in cui la scienza moderna superò le conquiste degli antichi. Fino ad allora, per tutto il corso del Medioevo, lo scopo era stato soltanto quello di comprendere l'opera degli antichi, ed ora, finalmente, furono risolte alcune questioni che gli antichi non erano riusciti a risolvere. E questo avvenne nel Cinquecento, cioè nel secolo prima dell'invenzione delle nuove branche della matematica: la geometria analitica, il calcolo differenziale e il calcolo integrale, che affermarono finalmente la superiorità della nuova scienza sulla vecchia. Dopo di ciò non ci fu

grande matematico che non tentasse di estendere le conquiste degli italiani e di risolvere in modo analogo per mezzo dei radicali equazioni di quinto, sesto e grado superiore.

L'eminente algebrista del XVII secolo, Tschirnhausen (1651-1708), credeva addirittura di aver finalmente trovato un metodo generale di soluzione. Il suo metodo era basato sulla trasformazione di un'equazione in una più semplice, ma proprio questa trasformazione richiedeva la soluzione di alcune equazioni ausiliarie. Successivamente, da un'analisi più approfondita si è dimostrato che il metodo di trasformazione di Tschirnhausen dà sì la soluzione delle equazioni di secondo, terzo e quarto grado, ma già per un'equazione di quinto grado richiede la soluzione preliminare di un'equazione ausiliaria di sesto grado, la cui soluzione a sua volta non era nota.

Fattorizzazione di un polinomio e formule di Viète.

Se accettiamo senza dimostrazione il cosiddetto teorema fondamentale dell'algebra[4] che ogni equazione

$$f(x) = 0,$$

dove $f(x) = x^n + a_1 x^{n-1} + \cdots + a_n$

è un polinomio in x di dato grado n e i coefficienti a_1, a_2, ⋯, a_n sono dati numeri reali o complessi, ha almeno una radice reale o complessa e teniamo conto che tutti i calcoli con numeri complessi vengono eseguiti seguendo le stesse regole dei numeri razionali, allora è facile dimostrare che il polinomio $f(x)$ può essere rappresentato (e in un solo modo) come prodotto di fattori di primo grado

$$f(x) = (x - a)(x - b) \cdots (x - l),$$

dove a, b, ⋯, l sono numeri reali o complessi.

Infatti, sia a una radice di $f(x)$; dividiamo $f(x)$ per $x - a$; poiché il divisore è di primo grado, il resto sarà un numero costante R, cioè avremo l'identità

$$f(x) = (x - a)\, f_1(x) + R,$$

dove $f_1(x)$ è un polinomio di grado $n - 1$ e R è una costante. Sostituendo qui al posto di x il numero a, otteniamo

$$f(a) = (a - a)\, f_1(a) + R = R.$$

Ma poiché a è una radice di $f(x)$, abbiamo $f(a) = 0$, e quindi $R = 0$, cioè un polinomio può sempre essere diviso per $(x - a)$ senza resto, dove a è una radice di questo polinomio. Così

$$f(x) = (x - a)\ f_1(x).$$

Ma se è vero il teorema fondamentale dell'algebra, allora a sua volta il polinomio $f_1(x)$ ha radice b, e otteniamo analogamente

$$f_1(x) = (x - b)\, f_2(x),$$

dove il polinomio $f_2(x)$ è già di grado (n - 2), ecc. Questa fattorizzazione, come si può facilmente dimostrare, è unica.

Ogni polinomio $f(x)$ di grado n ha in questo senso n e solo n radici a, b, ⋯, l. Queste radici possono essere tutte distinte ma può succedere che alcune di esse siano identiche. Allora diciamo che la radice corrispondente del polinomio $f(x)$ è una radice multipla con questa o quella molteplicità.

Moltiplicando l'espressione

$$(x - a)(x - b)(x - c) \cdots (x - l)$$

e confrontando i coefficienti delle stesse potenze di x, vediamo subito che

$$\begin{aligned} -a_1 &= a + b + c + \cdots + l, \\ a_2 &= ab + ac + \cdots + kl, \\ -a_3 &= abc + abd + \cdots, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \pm a_n &= abc \cdots l \end{aligned}$$

che sono le formule di Viète.

Un teorema sui polinomi simmetrici.

I membri delle formule di Viète sono polinomi nelle n lettere a, b, ⋯, l che non cambiano permutando queste lettere. Infatti, $a + b + \cdots + k + l = b + a + \cdots + k + l$, ecc. In generale, un polinomio in n lettere che rimane invariato per ogni loro permutazione si chiama polinomio simmetrico. Ad esempio, $5x^2 + 5y^2 - 7xy$ è simmetrico in x e y. Si può dimostrare che ogni polinomio simmetrico in n lettere, con coefficienti arbitrari A, B, ⋯, può essere espresso mediante sole addizioni, sottrazioni e moltiplicazioni in termini di questi coefficienti e dei polinomi di Viète nelle stesse lettere. Se a, b, ⋯, l sono le radici dell'equazione di grado n, $x^n + a_1 x^{n-1} + \cdots + a_n = 0$, allora ogni polinomio simmetrico nelle radici, con coefficienti A, B, ⋯, si può esprimere mediante le stesse operazioni in termini di A, B, ⋯ e dei coefficienti a_1, a_2, ⋯, a_n dell'equazione. Questo è il cosiddetto teorema fondamentale dei polinomi simmetrici.

I contributi di Lagrange.

Il famoso matematico francese Lagrange, nella sua grande opera “Riflessioni sulla soluzione delle equazioni algebriche” pubblicata nel 1770-1771 (con più di 200 pagine), esaminò criticamente tutte le soluzioni delle equazioni di secondo, terzo e quarto grado conosciute fino al suo tempo e dimostrò che il loro successo si basava sempre su proprietà che non valevano per equazioni di grado 5 o superiori. Dai tempi di del Ferro fino a quest'opera di Lagrange erano trascorsi più di due secoli e mezzo e nessuno durante questo lungo intervallo aveva dubitato della possibilità di risolvere equazioni di grado 5 e superiore per radicali, cioè di trovare formule comprendenti solo le operazioni di addizione, sottrazione, moltiplicazione, divisione e radicali con indici interi positivi, che esprimessero la soluzione di un'equazione in termini di coefficienti, cioè formule simili a quelle con cui era stata risolta l'equazione quadratica in antichità e le equazioni di terzo e quarto grado nel Cinquecento da parte degli italiani. Consideravano questa situazione dovuta solo alla loro incapacità di trovare una soluzione valida ma apparentemente profondamente nascosta.

Lagrange dice nelle sue memorie: "Il problema di risolvere (mediante radicali) equazioni di grado superiore a quattro è uno di quei problemi che non sono stati risolti sebbene nulla provi l'impossibilità di risolverli" e due pagine dopo completa questo: "Dal nostro ragionamento vediamo che è molto dubbio che i metodi che abbiamo considerato possano fornire una soluzione completa delle equazioni di quinto grado".

Nelle sue ricerche, Lagrange introdusse l'espressione

$$a + \epsilon b + \epsilon^2 c + \cdots + \epsilon^{n-1} l$$

nelle radici a, b, ⋯, l di un'equazione, dove ε è una n-esima radice dell'unità,[5] avendo stabilito che tali espressioni sono strettamente connesse con la soluzione delle equazioni per radicali. Queste espressioni sono oggi chiamate «risolventi di Lagrange».

Inoltre, Lagrange osservò che la teoria delle permutazioni delle radici di un'equazione è di grande importanza nella teoria della soluzione delle equazioni. Espresse addirittura il pensiero che la teoria delle permutazioni è la “vera filosofia dell'intera questione”, in cui aveva completamente ragione, come hanno dimostrato le successive ricerche di Galois.

Il metodo di soluzione delle equazioni di secondo, terzo e quarto grado di Lagrange non era lo stesso di quello italiano, che in ogni caso si basava su trasformazioni speciali di tipo complicato e per così dire accidentale. I metodi di Lagrange erano del tutto ordinati e si sviluppavano da un’idea generale che coinvolgeva la teoria dei polinomi simmetrici, la teoria delle permutazioni e la teoria dei risolventi.

Consideriamo, ad esempio, la soluzione con il metodo di Lagrange dell’equazione generale di quarto grado

$$x^4 + mx^3 + nx^2 + px + q = 0.$$

Siano le radici di questa equazione a, b, c, d. Consideriamo il risolvente

$$a + b - c - d,$$

cioè,

$$a + \epsilon c + \epsilon^2 b + \epsilon^3 d,$$

dove ϵ = - 1. Se permutiamo a, b, c, d in tutto 1·2·3·4 = 24 modi diversi, otteniamo complessivamente sei diverse espressioni

$$\begin{aligned} &a + b - c - d,\\ &a + c - b - d,\\ &a + d - c - b,\\ &c + d - a - b,\\ &b + d - a - c,\\ &b + c - a - d. \end{aligned} \qquad (5)$$

Un'equazione di sesto grado, le cui radici sono queste sei espressioni, avrà quindi coefficienti che non variano con tutte le 24 permutazioni di a, b, c, d, poiché ciascuna delle 24 permutazioni può permutare queste espressioni solo tra loro e i coefficienti dell'equazione di sesto grado non dipendono dall'ordine delle sue radici. Pertanto, questi coefficienti sono polinomi simmetrici in a, b, c, d. Ma poi, in virtù del teorema fondamentale sui polinomi simmetrici, questi coefficienti sono espressi mediante addizioni, sottrazioni e moltiplicazioni in termini dei coefficienti m, n, p, q dell'equazione. Inoltre, poiché le espressioni (5) sono a coppie di segni opposti, questa equazione di sesto grado conterrà solo termini di potenze pari. Infatti, se le espressioni (5) sono indicate rispettivamente con α, β, γ, - α, - β, - γ, allora il membro sinistro dell'equazione di sesto grado sarà uguale a

$$(y - \alpha)(y + \alpha)(y - \beta)(y + \beta)(y - \gamma)(y + \gamma) = (y^2 - \alpha^2)(y^2 - \beta)(y_2 - \gamma^2).$$

Il calcolo diretto fornisce l'equazione di sesto grado

$$y^6 - (3m^2 - 8n)y^4 + 3(m^4 - 16m^2n - 16n^2 + 16mp - 64q)y^2 - (m^2 - 4m + 8p)^2 = 0.$$

Prendendo $y^2 = t$, otteniamo un'equazione cubica in t, e se t', t'', t''' sono le sue radici, allora

$$\begin{aligned} a + b - c - d &= \sqrt{t'},\\ a + c - b - d &= \sqrt{t''},\\ a + d - b - c &= \sqrt{t'''}. \end{aligned}$$

Abbiamo anche

$$a + b + c + d = - m.$$

Sommando queste equazioni dopo la moltiplicazione per 1, 1, 1, 1 o 1, -1, -1, 1 o -1, 1, -1, 1 o -1, -1, 1, 1, otteniamo

$$a = \frac{1}{4}(-m + \sqrt{t'} + \sqrt{t''} + \sqrt{t'''}),$$

$$b = \frac{1}{4}(-m + \sqrt{t'} - \sqrt{t''} - \sqrt{t'''}),$$

$$c = \frac{1}{4}(-m - \sqrt{t'} + \sqrt{t''} - \sqrt{t'''}),$$

$$d = \frac{1}{4}(-m - \sqrt{t'} - \sqrt{t''} + \sqrt{t'''}).$$

Pertanto, la soluzione di un'equazione di quarto grado si riduce alla soluzione di un'equazione cubica; e le equazioni di terzo e secondo grado vengono risolte in modo analogo.

Lagrange ottenne grandi risultati nella teoria delle equazioni algebriche. Tuttavia, anche dopo i suoi persistenti sforzi, il problema della soluzione in radicali delle equazioni algebriche di grado superiore a 4 rimaneva da risolvere. Questo problema, sul quale i matematici avevano lavorato invano per quasi tre secoli, costituiva, secondo l'espressione di Lagrange, "una sfida alla mente umana".

La scoperta di Abel.

Di conseguenza fu una grande sorpresa per tutti i matematici quando nel 1824 venne alla luce il lavoro di un giovane genio norvegese Abel (1802-1829), in cui veniva fornita una dimostrazione che se i coefficienti di un'equazione a_1, a_2, ···, a_n, sono considerati semplicemente come lettere, allora non esiste alcuna espressione radicale in questi coefficienti che sia radice dell'equazione corrispondente, se il suo grado $n \geqslant 5$. Pertanto, per tre secoli, gli sforzi dei più grandi matematici di tutti i paesi per risolvere equazioni di grado 5 o superiore per radicali non hanno portato al successo per la semplice ragione che questo problema semplicemente non ha soluzione.

Tale formula è nota per le equazioni di secondo grado e, come abbiamo visto, esistono formule analoghe per le equazioni di terzo e quarto grado, ma per le equazioni di grado 5 o superiore non esistono formule del genere.

La dimostrazione di Abel è difficile e non la forniremo qui.

Teoria di Galois.

Ma questo non era ancora tutto. Un risultato davvero notevole nella teoria delle equazioni algebriche doveva ancora venire. Il punto è che ci sono arbitrariamente molte forme speciali di equazioni di qualsiasi grado che sono risolvibili in radicali, e molte di esse sono esattamente quelle equazioni che sono importanti nelle applicazioni. Tali, ad esempio, sono le equazioni binomiali $x^n = A$. Abel trovò un'altra classe molto ampia di tali equazioni, le cosiddette equazioni cicliche e le ancora più generali equazioni "Abeliane". In relazione al problema della costruzione con riga e compasso di poligoni regolari, Gauss considerò esplicitamente le cosiddette equazioni ciclotomiche, cioè equazioni della forma

$$x^{p-1} + x^{p-2} + \cdots + x + 1 = 0,$$

dove p è un numero primo, e dimostrò che esse possono sempre essere ridotti ad una catena di equazioni di grado inferiore; inoltre, trovò le condizioni necessarie e sufficienti affinché tale equazione possa essere risolta in radici quadrate. La necessità di queste condizioni fu rigorosamente dimostrata solo da Galois.

Pertanto, dopo il lavoro di Abel la situazione era la seguente: sebbene, come mostrato da Abel, l'equazione generale di grado superiore a 4 non possa essere risolta mediante radicali, esistono arbitrariamente molte diverse equazioni speciali di grado arbitrario, che possono tutte essere risolte mediante radicali. L'intera questione della risoluzione delle equazioni in radicali fu posta da queste scoperte su un terreno completamente nuovo. Divenne chiaro che il compito ora era determinare esattamente quali equazioni possono essere risolte con radicali, o in altre parole, quali sono le condizioni necessarie e sufficienti per la risolubilità di un'equazione in radicali. Questo problema, la cui risposta diede in un certo senso la delucidazione finale dell'intero problema, fu risolto dall'geniale matematico francese Évariste Galois.

Galois (1811-1832) morì all'età di 20 anni in un duello. Negli ultimi due anni della sua vita non poté dedicare molto tempo alla matematica, poiché fu trascinato dal turbinio tempestoso della vita politica al tempo della Rivoluzione del 1830 e languì in prigione per il suo discorso contro il regime reazionario di Luigi Filippo. Tuttavia, nella sua breve vita, Galois fece scoperte molto in anticipo sui suoi tempi in vari settori della matematica e in particolare produsse risultati davvero notevoli nella teoria delle equazioni algebriche. In una piccola pubblicazione "Memorie sulle condizioni di risolubilità delle equazioni in radicali", rimasta manoscritta dopo la sua morte e pubblicata per la prima volta nel 1846 da Liouville, Galois partì da alcuni concetti molto semplici ma profondi e finalmente districò l'intero complesso di difficoltà relative alla soluzione delle equazioni in radicali, difficoltà con le quali fino ai suoi tempi avevano lottato senza successo i matematici più illustri. Il successo di Galois si basava sul fatto che per la prima volta introdusse nella teoria delle equazioni una serie di nuovi concetti generali estremamente importanti, che successivamente giocarono un ruolo importante nella matematica nel suo insieme.

Consideriamo la teoria di Galois per un caso speciale, vale a dire quando i coefficienti a_1, a_2, ⋯, a_n dell'equazione di grado n data

$$x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0 \qquad (6)$$

sono numeri razionali. Questo caso è particolarmente interessante e comporta già sostanzialmente tutte le difficoltà della teoria generale di Galois. Assumeremo anche che le radici a, b, c, ⋯ di questa equazione siano distinte.

Galois inizia, come Lagrange, considerando un'espressione di primo grado in a, b, c, ⋯

$$V = Aa + Bb + Cc + \cdots,$$

sebbene non richieda che i coefficienti A, B, C, ⋯ di questa espressione siano radici dell'unità, ma prende per A, B, C, ⋯ qualsiasi numero intero tale da dare valori numericamente distinti per tutte le $n! = 1{\cdot}2{\cdot}3{\cdots}n$ quantità V, V', V'', ⋯, $V^{(n!-1)}$ ottenute da V permutando le radici a, b, c, ⋯ in tutto $n!$ modi possibili. Questo può sempre essere fatto. Quindi Galois costruisce l'equazione di grado $n!$ le cui radici sono V, V', V'', ⋯, $V^{(n!-1)}$. Il teorema sui polinomi simmetrici mostra che i coefficienti di questa equazione $\Phi(x) = 0$ di grado $n!$ saranno numeri razionali.

Finora tutto è abbastanza simile a quanto ha fatto Lagrange.

Successivamente Galois introdusse il primo importante nuovo concetto, il concetto di irriducibilità di un polinomio in un dato campo di numeri. Se viene fornito un polinomio in x, i cui coefficienti, ad esempio, sono numeri razionali, allora il polinomio si dice riducibile nel campo dei numeri razionali se può essere rappresentato sotto forma di un prodotto di polinomi di grado inferiore con coefficienti razionali. In caso contrario, nel campo dei numeri razionali il polinomio si dice irriducibile. Il polinomio $x^3 - x^2 - 4x - 6$ è riducibile nel campo dei numeri razionali, poiché è uguale a $(x^2 + 2x + 2)(x - 3)$, ma ad esempio il polinomio $x^3 + 3x^2 + 3x - 5$ è irriducibile, come si può vedere, nel campo dei numeri razionali.

Esistono metodi, che richiedono certamente lunghi calcoli, per fattorizzare qualsiasi dato polinomio con coefficienti razionali in polinomi irriducibili nel campo dei numeri razionali.

Galois quindi fattorizza il polinomio $\Phi(x)$ di cui sopra in fattori irriducibili nel campo dei numeri razionali.

Sia $F(x)$ uno di questi polinomi irriducibili (quale di essi è irrilevante per quanto segue) e sia di grado m.

Il polinomio $F(x)$ sarà quindi il prodotto di m degli $n!$ fattori di primo grado $x - V$, $x - V'$, ⋯, $x - V^{(n!-1)}$, nei quali il polinomio di grado $n!$ $\Phi(x)$ è stato scomposto. Siano questi m fattori $x - V$, $x - V'$, ⋯, $x - V^{(m-1)}$. Enumeriamo in qualsiasi ordine le radici, a, b, c, ⋯, l della data equazione di grado n (6) dando loro gli indici, 1,2, ⋯, n. Allora le quantità V, V', ⋯, $V^{(n!-1)}$ corrispondono a tutte le $n!$ permutazioni possibili dei numeri 1, 2, ⋯, n, corrispondenti alle permutazioni delle radici, e V, V', ⋯, $V^{(m-1)}$ corrispondono solo a m di queste permutazioni. L'insieme G di queste m permutazioni dei numeri 1, 2, ⋯, n è chiamato gruppo di Galois dell'equazione data (6).[6]

Galois introduce poi alcuni nuovi concetti e sviluppa argomentazioni semplici ma davvero notevoli con le quali dimostra che una condizione necessaria e sufficiente per la risolubilità

dell'equazione (6) in radicali è che il gruppo *G* di permutazioni dei numeri 1, 2, ⋯, *n* soddisfi una certa condizione definita.

Pertanto, la profezia di Lagrange secondo cui alla base dell'intero problema c'era la teoria delle permutazioni si è rivelata vera.

In particolare, il teorema di Abel sulla non risolubilità di un'equazione generale di quinto grado in radicali può ora essere dimostrato come segue. Si può dimostrare che esistono arbitrariamente molte equazioni di quinto grado, anche con coefficienti interi per le quali il corrispondente polinomio di 120° grado $\Phi(x)$ è irriducibile, cioè il cui gruppo di Galois è il gruppo di tutte le 5! = 120 permutazioni degli indici 1, 2, 3, 4, 5 delle sue radici. Ma questo gruppo, come si può mostrare, non soddisfa il criterio di Galois, e quindi queste equazioni di quinto grado non possono essere risolte per radicali.

Ad esempio, si può dimostrare che l'equazione $x^5 + x - a = 0$, dove a è un numero intero positivo, nella maggior parte dei casi non può essere risolta mediante radicali. Ad esempio, non è risolvibile in radicali per $a = 3,4, 5, 7, 8,9, 10, 11, \cdots$

Applicazione della teoria di Galois al problema della risolubilità di problemi geometrici mediante riga e compasso.

Una delle applicazioni speciali più notevoli della teoria di Galois è la seguente. Molti problemi di geometria piana possono essere risolti soltanto con la costruzione di riga e compasso. Ad esempio, possiamo costruire con riga e compasso un triangolo regolare, un quadrato, un pentagono, un esagono, un ottagono, un decagono, ecc., ma è impossibile costruire un poligono regolare di sette, nove o undici lati. Quali problemi possono essere risolti con riga e compasso e quali no? Prima di Galois era un problema irrisolto. Dalla teoria di Galois otteniamo la seguente risposta.

La soluzione simultanea di equazioni di due rette, una retta e un cerchio, o due cerchi, può essere ridotta alla soluzione di equazioni di primo o secondo grado. Per una retta e un cerchio è chiaro, e nel caso di due cerchi $(x - a_1)^2 + (y - b_1)^2 = r_1^2$ e $(x - a_2)^2 + (y - b_2)^2 = r_2^2$ se sottraiamo un'equazione dall'altra, le x^2 e y^2 si annullano, e otteniamo un'equazione di primo grado, che va risolta contemporaneamente con l'equazione di uno dei cerchi, così che ancora una volta abbiamo un'equazione quadratica. Pertanto, ogni passo del problema da risolvere con riga e compasso si riduce ad un'equazione di primo o secondo grado, e di conseguenza tutti i problemi risolvibili con riga e compasso si riducono ad un'equazione algebrica ad una incognita, la cui soluzione prevede l'estrazione di una catena di radici quadrate. Al contrario, se la soluzione di un problema geometrico si riduce a una simile equazione algebrica, allora può essere risolta con riga e compasso, poiché le radici quadrate, come è noto, possono essere costruite con riga e compasso.

Se è dato un problema geometrico, dobbiamo prima stabilire un'equazione algebrica equivalente al problema dato. Se è impossibile stabilire una simile equazione, il problema ovviamente non è risolvibile con riga e compasso. Se l'equazione è stata stabilita, allora dobbiamo selezionare uno dei suoi fattori irriducibili che è connesso con la soluzione del problema e determinare se questa equazione irriducibile può essere risolta in radici quadrate.

Come mostra la teoria di Galois, per questo è necessario e sufficiente che il numero *m* di permutazioni che ne costituiscono il gruppo di Galois sia una potenza di 2.

Con questo test possiamo dimostrare il teorema enunciato da Gauss secondo cui un poligono regolare con numero primo *p* di lati può essere costruito con riga e compasso se e solo se il numero primo *p* ha la forma $2^{2^n} + 1$, cioè per *p* = 3, 5, 17, 257 ma non per *p* = 7, 11, 13, 19, 23, 29, 31, ⋯, ecc. Gauss dimostrò solo la parte "se" di questa affermazione.

Con lo stesso metodo possiamo dimostrare che è impossibile dividere un angolo arbitrario in tre parti uguali con riga e compasso, o duplicare il cubo, cioè dallo spigolo di un dato cubo trovare lo spigolo di un cubo con un volume doppio, e così via.

L'impossibilità di quadrare il cerchio, cioè di costruire con riga e compasso il lato di un quadrato uguale in area a un cerchio di raggio dato, si dimostra in modo diverso. Si può cioè dimostrare che il lato di un tale quadrato non è collegato al raggio da alcuna equazione algebrica, cioè è per così dire trascendente rispetto al raggio, e di conseguenza non è a fortiori esprimibile in termini di raggio mediante una catena di radici quadrate. Questa dimostrazione è difficile e non segue dalla teoria di Galois.

Due problemi fondamentali irrisolti legati alla teoria di Galois.

Nella teoria di Galois rimangono altri due problemi fondamentali che non sono stati ancora risolti nella loro forma generale, sebbene molti eccellenti matematici vi abbiano lavorato quasi ininterrottamente.

Il primo di questi è il problema dei cosiddetti risolventi di Hilbert-Čebotarev (da non confondere con i risolventi di Lagrange) che è una generalizzazione diretta del problema della soluzione delle equazioni in radicali. L'idea è questa: dire che un'equazione è risolvibile in radicali è esattamente come dire che la sua soluzione si riduce a una catena di equazioni binomiali successive, poiché il radicale $\sqrt[n]{A}$ è una radice dell'equazione binomiale $x^n = A$. Ma può succedere che, sebbene l'equazione non possa essere ridotta a una catena di equazioni semplici come quelle binomiali, possa tuttavia essere ridotta a una catena di alcune altre equazioni molto semplici. Già alla fine del XVIII secolo era stato dimostrato che l'equazione generale di quinto grado può essere ridotta ad una catena di equazioni binomiali insieme ad un'ulteriore equazione della forma $x^5 + x + A = 0$, la quale, pur non essendo binomiale, ha come le equazioni binomiali un solo parametro *A*.

Successivamente si dimostrò che anche un'equazione di sesto grado non può essere ridotta a una catena di equazioni ad un parametro. Per le equazioni di qualsiasi grado dobbiamo risolvere il problema: che tipo di equazioni più semplici, cioè con un numero minimo di parametri, compongono la catena a cui può essere ridotta la nostra equazione.

Se l'equazione data viene ridotta a una catena di equazioni a un parametro di tipo definito, allora per ciascuna di queste equazioni a un parametro possiamo calcolare una tabella, fornendo le sue radici in funzione del suo parametro. Quindi la soluzione dell'equazione data si riduce all'uso di una catena di tali tabelle.

In secondo luogo, un problema ancora più profondo consiste nell'inverso della teoria di Galois. Galois dimostrò che le proprietà delle soluzioni di un'equazione dipendono dal suo gruppo. Ma viceversa, può un qualsiasi gruppo di permutazioni essere il gruppo di Galois di qualche equazione e possiamo impostare tutte le equazioni il cui gruppo di Galois è un dato gruppo?

Per quanto riguarda la prima di queste due domande si conoscono solo risultati parziali, sebbene matematici eccezionali come Klein e Hilbert vi abbiano lavorato con tenacia; i primi teoremi generali furono formulati dallo straordinario algebrista sovietico N. G. Čebotarev.

La seconda questione per i cosiddetti gruppi risolvibili, cioè i gruppi che soddisfano il criterio di Galois, è stata risolta affermativamente negli ultimi anni dal matematico sovietico I. R. Šafarevič.

§ 3. Il teorema fondamentale dell'algebra

Nella sezione precedente abbiamo considerato i tentativi, durati tre secoli, di risolvere per radicali un'equazione di $n°$ grado. Il problema si rivelò molto profondo e difficile e portò alla creazione di nuovi concetti, importanti non solo per l'algebra ma anche per la matematica nel suo insieme. Per quanto riguarda la soluzione pratica delle equazioni, il risultato di tutto questo lavoro fu il seguente. È diventato chiaro che la soluzione per radicali è lungi dall'essere disponibile per tutte le equazioni algebriche e, anche quando è disponibile, ha poco valore pratico a causa della sua complessità, tranne nel caso dell'equazione quadratica.

In considerazione di ciò, i matematici hanno cominciato molto tempo fa a lavorare sulla teoria delle equazioni algebriche in tre direzioni completamente diverse, vale a dire: (1) sul problema dell'esistenza di una radice; (2) sul problema di come possiamo imparare dai coefficienti dell'equazione qualcosa sulle sue radici senza risolverla; ad esempio, ha radici reali e quante; e infine (3) sul calcolo approssimato delle radici di un'equazione.

Innanzitutto era necessario dimostrare che in generale qualsiasi equazione algebrica di $n°$ grado a coefficienti reali o complessi ha sempre almeno una radice reale o complessa.[7]

Questo teorema, che è uno dei più importanti di tutta la matematica, rimase per molto tempo senza una dimostrazione rigorosa. In considerazione della sua importanza e difficoltà, è generalmente chiamato il "teorema fondamentale dell'algebra", sebbene la maggior parte dei metodi con cui è stato dimostrato siano strettamente legati all'analisi infinitesimale quanto all'algebra. La prima dimostrazione fu data da d'Alembert. Un punto della dimostrazione di d'Alembert, come fu chiarito in seguito, si rivelò difettoso. Vale a dire, d'Alembert assunse come banale la proposizione generale dell'analisi secondo cui una funzione continua, data su un insieme limitato e chiuso di punti, ha in qualche punto dell'insieme un minimo. Questo è vero ma andava dimostrato. Una prova rigorosa di questa proprietà si ottenne solo nella seconda metà dell'Ottocento, cioè cento anni dopo le indagini di d'Alembert.

Si ritiene generalmente che la prima dimostrazione rigorosa del teorema fondamentale dell'algebra sia stata data da Gauss; tuttavia, alcune delle sue dimostrazioni richiedono, per essere pienamente rigorose, aggiunte non minori di quelle richieste per la dimostrazione di d'Alembert. Oggi sono note diverse dimostrazioni completamente rigorose di questo teorema.

Nella presente sezione consideriamo la dimostrazione del teorema fondamentale dell'algebra basata sul cosiddetto lemma di d'Alembert, e diamo anche una dimostrazione completa della suddetta proposizione dell'analisi.

La teoria dei numeri complessi.

Prima di considerare la dimostrazione del teorema fondamentale dell'algebra, dobbiamo innanzitutto richiamare la teoria dei numeri complessi studiata alle scuole superiori. Le difficoltà che hanno portato alla creazione della teoria dei numeri complessi si incontrano per la prima volta nella risoluzione delle equazioni quadratiche. Cosa dobbiamo fare se il numero $p^2/4 - q$ sotto la radice quadrata nella formula per la soluzione dell'equazione quadratica è negativo? Non esiste un numero reale, positivo o negativo, che sia la radice quadrata di un numero negativo, poiché il quadrato di qualsiasi numero reale è positivo o zero.

Dopo lunghi dubbi, durati più di un secolo, i matematici sono giunti alla conclusione che è necessario introdurre una nuova forma di numeri, i cosiddetti numeri complessi, con le seguenti leggi di operazioni su di essi.

Convenzionalmente viene introdotto un numero di natura nuova: $i = \sqrt{-1}$ tale che $i^2 = -1$, e vengono considerati i numeri della forma $a + bi$, dove a e b sono numeri reali ordinari. I numeri $a + bi$ sono detti complessi. Due di questi numeri $a + bi$ e $c + di$ sono considerati uguali se $a = c$, $b = d$. La somma di due di questi numeri è definita come il numero $(a + c) + (b + d)i$, e la loro differenza è il numero $(a - c) + (b - d)i$. Nella moltiplicazione concordiamo di moltiplicare questi numeri come se fossero binomi ma di tenere in considerazione che $i^2 = -1$, ovvero,

$$(a + bi)(c + di) = ac + bci + adi + bdi^2 = (ac - bd) + (bc + ad)i.$$

Se a e b sono considerati come coordinate rettangolari di un punto e il punto è associato al numero complesso $a + bi$, allora l'addizione e la sottrazione di numeri complessi corrispondono all'addizione e alla sottrazione di vettori, cioè di segmenti diretti dall'origine ai punti con coordinate (a, b) e (c, d), poiché nell'addizione dei vettori si sommano le coordinate corrispondenti. Quanto al significato geometrico di un prodotto nel cosiddetto piano dei numeri complessi, lo possiamo vedere più facilmente se consideriamo la lunghezza ρ del vettore dall'origine del sistema di coordinate al punto (x, y) (tale lunghezza è chiamata modulo del numero complesso $z = x + iy$) e l'angolo ϕ che il vettore forma con l'asse Ox (questo angolo si chiama argomento del numero complesso $z = x + iy$); in altre parole, se consideriamo non le coordinate cartesiane x e y ma le cosiddette coordinate polari ρ e ϕ (figura 1). Allora $x = \rho \cos \phi$, $y = \rho \sin \phi$ e di conseguenza il numero complesso stesso può essere scritto come

$$x + iy = \rho(\cos\phi + i\sin\phi).$$

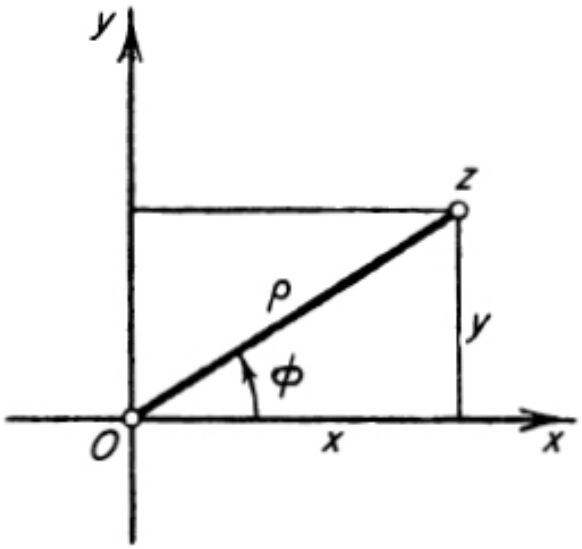


Figura 1.

Se

$$a + bi = \rho_1(\cos\phi_1 + i\sin\phi_1), \quad c + di = \rho_2(\cos\phi_2 + i\sin\phi_2),$$

allora

$$\begin{aligned} ac - bd &= \rho_1\rho_2(\cos\phi_1\cos\phi_2 - \sin\phi_1\sin\phi_2) = \rho_1\rho_2\cos(\phi_1 + \phi_2), \\ bc + ad &= \rho_1\rho_2(\sin\phi_1\cos\phi_2 + \cos\phi_1\sin\phi_2) = \rho_1\rho_2\sin(\phi_1 + \phi_2), \end{aligned}$$

da ciò vediamo che nella moltiplicazione di due numeri complessi si moltiplicano i loro moduli ρ_1 e ρ_2 e si sommano gli argomenti ϕ_1 e ϕ_2. Nella divisione, essendo l'operazione inversa della moltiplicazione, si divide un modulo per l'altro e si sottraggono gli argomenti

$$\begin{aligned} &\rho_1(\cos\phi_1 + i\sin\phi_1)\,\rho_2(\cos\phi_2 + i\sin\phi_2) \\ &\qquad = \rho_1\rho_2[\cos(\phi_1 + \phi_2) + i\sin(\phi_1 + \phi_2)] \end{aligned}$$

E

$$\frac{\rho_1(\cos\phi_1 + i\sin\phi_1)}{\rho_2(\cos\phi_2 + i\sin\phi_2)} = \frac{\rho_1}{\rho_2}[\cos(\phi_1 - \phi_2) + i\sin(\phi_1 - \phi_2)].$$

Elevando a una potenza con esponente intero positivo n, di conseguenza, il modulo viene elevato alla stessa potenza n-esima e l'argomento viene moltiplicato per n

$$[\rho(\cos\phi + i\sin\phi)]^n = \rho^n(\cos n\phi + i\sin n\phi).$$

Viceversa, estraendo le radici

$$\sqrt[n]{\rho(\cos\phi + i\sin\phi)} = \sqrt[n]{\rho}\left(\cos\frac{\phi}{n} + i\sin\frac{\phi}{n}\right).$$

Tuttavia, nell'estrazione delle radici si presenta una situazione speciale. Sia n un esponente intero positivo. Allora

$$\sqrt[n]{\rho(\cos\phi + i\sin\phi)}$$

è uguale al numero

$$\sqrt[n]{\rho}\left(\cos\frac{\phi}{n} + i\sin\frac{\phi}{n}\right)$$

poiché elevando questo numero alla potenza n si ottiene il radicando.

Ma questo è solo un valore della radice. Il punto è che il numero complesso

$$\sqrt[n]{\rho}\left[\cos\left(\frac{\phi}{n} + \frac{2k\pi}{n}\right) + i\sin\left(\frac{\phi}{n} + \frac{2k\pi}{n}\right)\right],$$

dove k è uno qualsiasi dei numeri 1, 2, ⋯, n - 1, sarà anche una n-esima radice del numero

$$\rho(\cos\phi + i\sin\phi).$$

Infatti, secondo la regola per l'elevazione a una potenza, se eleviamo questo numero alla potenza n-esima, otteniamo il numero

$$(\sqrt[n]{\rho})^n\left[\cos n\left(\frac{\phi}{n} + \frac{2k\pi}{n}\right) + i\sin n\left(\frac{\phi}{n} + \frac{2k\pi}{n}\right)\right]$$
$$= \rho[\cos(\phi + 2k\pi) + i\sin(\phi + 2k\pi)],$$

dove l'addendo $2k\pi$, a causa delle proprietà di seno e coseno, può essere trascurato, poiché non cambia né seno né coseno. Quindi anche la n-esima potenza di questo numero è

$$\rho(\cos\phi + i\sin\phi),$$

cioè, questo numero è

$$\sqrt[n]{\rho(\cos\phi + i\sin\phi)}.$$

È facile vedere che nessun altro numero complesso, oltre a questi n numeri per $k = 0, 1, 2, \cdots, n - 1$ è una radice n-esima di

$$\rho(\cos\phi + i\sin\phi).$$

Geometricamente, l'estrazione delle radici n-esime si descrive così. I punti del piano complesso che rappresentano i valori della radice $\sqrt[n]{}$ del numero $\rho(\cos\phi + i\sin\phi)$ sono i

vertici di un poligono regolare di n lati, inscritto nella circonferenza con centro nell'origine e raggio $\sqrt[n]{\rho}$, orientato in modo che uno dei vertici abbia argomento ϕ/n (figura 2).

Facciamo la seguente osservazione. Se

$$f(z) = z^n + c_1 z^{n-1} + \cdots + c_{n-1} z + c_n$$

è un polinomio in z con coefficienti reali o complessi dati c_1, c_2, ⋯, c_n, e cambiamo z continuamente, cioè spostiamo continuamente il punto $z = x + iy$ nel piano complesso, allora anche il punto complesso $Z = X + iY = f(z)$ si sposterà continuamente nel piano complesso. Ciò risulta evidente dal fatto che se sostituiamo in $f(z)$ il valore di $z = x + iy$, $c_1 = a_1 + b_1 i$, $c_2 = a_2 + b_2 i$, ⋯, $c_n = a_n + b_n i$ ed eseguiamo tutti i calcoli, troviamo che

$$f(z) = X + iY,$$

Dove

$$X = P(x, y), \quad Y = Q(x, y)$$

sono polinomi di grado n in x e y con coefficienti reali espressi in termini di a_i e b_i. In caso di cambiamento continuo di x e y, anche questi polinomi cambieranno continuamente.

Notiamo inoltre che, poiché il modulo $\rho = |f(z)|$ è uguale a $\sqrt{X^2 + Y^2}$, durante uno spostamento continuo del punto z nel piano complesso, il modulo $|f(z)|$ cambierà anche continuamente. In altre parole, se il punto z è sufficientemente vicino al punto α, allora la differenza $|f(z)| - |f(\alpha)|$ dei valori assoluti è inferiore a qualsiasi numero positivo preassegnato.

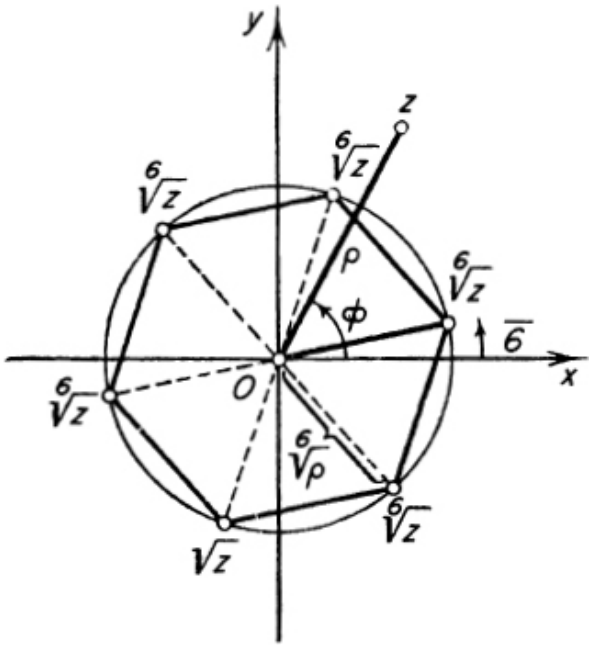


Figura 2.

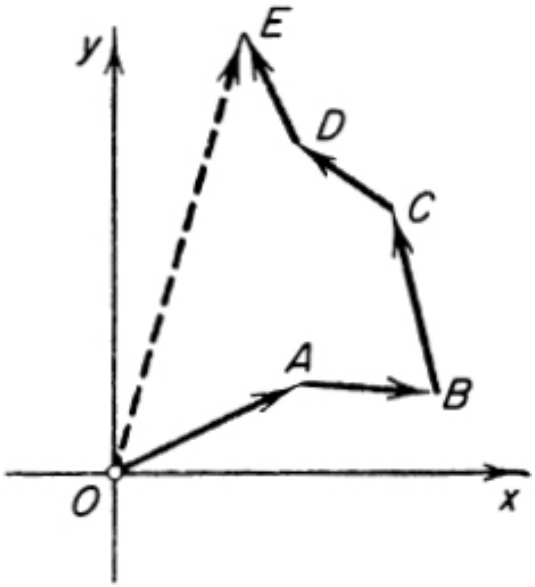


Figura 3.

Notiamo anche che il modulo di una somma di più numeri complessi è sempre minore o uguale alla somma dei moduli di questi numeri, il che equivale a dire che il segmento rettilineo OE (figura 3) è più corto o uguale alla linea poligonale OABCDE, essendo ad essa uguale se e solo se tutti i suoi segmenti giacciono su una stessa retta e in una sola direzione.

Ricordiamo infine che dire “un numero complesso è uguale a zero” equivale a dire che “il suo modulo è uguale a zero”, poiché il modulo ρ di un numero complesso è la distanza dall'origine al punto corrispondente.

Applichiamo ora la teoria dei numeri complessi alla dimostrazione del teorema fondamentale dell'algebra; anche se va notato che il significato della teoria dei numeri complessi va ben oltre i limiti dell'algebra. In molti ambiti della matematica diversi dall’algebra, non possiamo farne a meno. In molte applicazioni, ad esempio nella teoria delle correnti alternate, numerosi problemi vengono risolti nel modo più semplice mediante numeri complessi. Ma ciò che è più importante è l'applicazione dei numeri complessi, o più precisamente della teoria delle funzioni di una variabile complessa, alla teoria di certe funzioni speciali di due variabili reali che si chiamano armoniche. Per mezzo di queste funzioni si possono risolvere importanti problemi della teoria del volo aereo, della conduzione del calore in una piastra, dei campi elettrici piani e dell'elasticità. Il fondatore dell'aerodinamica contemporanea, N. E. Žukovskij, ha ricavato il famoso teorema sulla forza portante sull'ala di un aereo, attraverso lo studio delle funzioni di una variabile complessa.[8]

Passiamo ora alla dimostrazione del teorema fondamentale dell'algebra.

Teorema.

Qualsiasi polinomio

$$f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n ,$$

i cui coefficienti

$$a_0 , a_1 , \cdots, a_{n-1} , a_n$$

sono numeri arbitrari reali o complessi, ha almeno una radice reale o complessa.

Assumeremo che il polinomio dato sia di grado *n*, cioè che $a_0 \neq 0$.

La superficie del modulo di un polinomio.

Consideriamo l'intero problema geometricamente. Sopra ogni punto *z* del piano complesso erigiamo un segmento perpendicolare *t*, di lunghezza uguale al modulo | *f*(*z*) | del polinomio *f*(*z*) in questo punto *z*. Gli estremi di queste quote definiscono una superficie *M*, che può essere chiamata superficie del modulo del polinomio *f*(*z*). Vediamo che questa superficie: (1) non scende mai al di sotto del piano complesso, poiché il modulo di qualsiasi numero complesso (in questo caso, il numero *f*(*z*)) non è negativo; (2) per ogni dato punto *z* del piano complesso, la superficie ha uno ed un solo punto che o giace verticalmente sopra tale punto oppure coincide con esso, cioè la superficie *M* si estende in un foglio sopra tutto il piano complesso e può in alcuni punti toccare il piano stesso; (3) la superficie è continua nel senso che una variazione continua della posizione del punto *z* sul piano complesso produce una variazione continua del valore di *t* = | *f*(*z*) |, cioè nella quota *t* di punti della superficie. (Questo è stato mostrato nell'ultima sottosezione.)

Il teorema fondamentale dell'algebra consiste nel dimostrare che la superficie *M* tocca il piano complesso almeno in un punto e non rimane ovunque a distanza positiva sopra di esso.

Sulla crescita del modulo di un polinomio al crescere della distanza dall'origine.

Mostriamo che non importa quanto sia grande il numero positivo *G*, possiamo trovare un raggio *R* tale che per tutti i punti *z* del piano complesso, che si trovano all'esterno del cerchio di raggio *R* con centro nell'origine, l'altezza *t* dei punti della superficie *M* sopra il piano complesso è maggiore di *G*.

Scriviamo infatti il polinomio *f*(*z*) come

$$a_0 z^n \left[1 + \left(\frac{a_1}{a_0 z} + \frac{a_2}{a_0 z^2} + \cdots + \frac{a_n}{a_0 z^n}\right)\right].$$

Il modulo dell'espressione

$$\left(\frac{a_1}{a_0 z} + \frac{a_2}{a_0 z^2} + \cdots + \frac{a_n}{a_0 z^n}\right)$$

non è maggiore della somma dei moduli degli addendi

$$\left|\frac{a_1}{a_0 z}\right| + \left|\frac{a_2}{a_0 z^2}\right| + \cdots + \left|\frac{a_n}{a_0 z^n}\right|,$$

e, all'aumentare del modulo di *z*, ognuno di questi addendi diminuisce, per cui diminuisce anche la somma. Pertanto, per tutti gli *z* i cui moduli sono maggiori di un certo numero *R′*, il modulo di questa espressione tra parentesi è inferiore, ad esempio, a $\frac{1}{2}$.

Ma poi per tutti questi *z*, l'espressione

$$\Omega = \left[1 + \left(\frac{a_1}{a_0 z} + \frac{a_2}{a_0 z^2} + \cdots + \frac{a_n}{a_0 z^n}\right)\right]$$

avrà un modulo maggiore di $\frac{1}{2}$. Il modulo del primo fattore $a_0 z^n$ è pari a $| a_0 | \cdot | z |^n$, quindi aumenta al crescere del modulo di *z*; inoltre, aumenta oltre ogni limite. Pertanto, non importa quanto sia grande il numero positivo *G*, esiste un numero positivo *R* tale che per tutti gli *z*, i cui moduli sono maggiori di *R*, $| f(z) | = | a_0 | \cdot | z |^n \cdot | \Omega |$ è maggiore di *G*.

L'esistenza dei minimi della superficie *M*. Diremo che in un punto *α* del piano complesso la superficie *M* ha un minimo se il valore dell'altezza *t* del punto della superficie *M* sopra questo punto *α* è minore o uguale ai suoi valori in tutti i punti di qualche intorno del punto *α*, cioè in tutti i punti di un cerchio, per quanto piccola, con centro nel punto *α*.

Sia l'altezza *t* del punto della superficie *M* corrispondente all'origine, cioè al punto *z* = 0 del piano complesso, uguale a *g*, cioè $|f(0)| = g$. Prendiamo $G > g$. Tutte le altezze *t* dei punti della superficie *M* non sono negative e cambiano continuamente durante il movimento continuo del punto *z* nel piano complesso. La superficie *M* ha altezza $t > G$ all'esterno di un cerchio disegnato attorno all'origine con raggio *R* e altezza $t = g < G$ al centro del cerchio. D'Alembert considerava una conseguenza ovvia che da qualche parte all'interno del cerchio *R* esiste un punto in cui l'altezza è minima; più precisamente, dove il valore di *t* è minore o uguale ai suoi valori in tutti i rimanenti punti del cerchio *R*, cioè la superficie *M* ha almeno un minimo.

La prova rigorosa dell'esistenza di tale minimo si basa sul seguente assioma di continuità dell'insieme dei numeri reali.

Se vengono date due successioni di numeri reali: $a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n \cdots$ e $b_1 \geqslant b_2 \geqslant \cdots \geqslant b_n \geqslant \cdots$, tali che $b_n > a_n$, per tutti *n* e $b_n - a_n \to 0$ per $n \to \infty$, allora esiste uno e un solo numero reale *c*, tale che $a_n \leqslant c \leqslant b_n$ per tutti gli *n*.

Geometricamente, questa proprietà di continuità significa che se sulla linea è data una successione di intervalli $[a_n, b_n]$ (figura 4), tale che ogni intervallo successivo è contenuto nell'intervallo precedente, e le lunghezze degli intervalli diventano arbitrariamente piccole, allora esiste un punto *c* appartenente a tutti gli intervalli della successione. In altre parole, gli intervalli "si riducono" fino a diventare un punto e non "un posto vuoto".

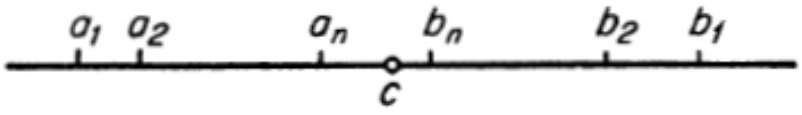


Figura 4.

Poiché la lunghezza del segmento $[a_n, b_n]$ si avvicina a zero all'aumentare di *n*, esiste un solo punto *c*. Dalla proprietà di continuità per l'insieme di tutti i punti sull'asse dei numeri segue

immediatamente la proprietà di continuità per i numeri complessi, cioè per i punti del piano. Diamo una formulazione geometrica di questa proprietà.

Se nel piano è data una successione di rettangoli $\Delta_1, \Delta_2, \cdots, \Delta_n, \cdots$, con lati paralleli agli assi coordinati, tale che ogni rettangolo è contenuto nel precedente, e tale che la lunghezza delle loro diagonali diminuisce indefinitamente, allora esiste uno ed un solo punto che è contenuto in tutti i rettangoli della successione. Questa proprietà di continuità del piano deriva direttamente dalla proprietà di continuità della linea. Per la dimostrazione è sufficiente proiettare i rettangoli sugli assi coordinati.

Ora è facile stabilire il cosiddetto teorema di Bolzano-Weierstrass.

Se in un rettangolo è data una successione infinita di punti $z_1, z_2, \cdots, z_n, \cdots$, allora all'interno o sul confine del rettangolo esiste un punto z_0, tale che in qualsiasi intorno arbitrariamente piccolo di z_0, cioè all'interno di un cerchio arbitrariamente piccolo con centro in z_0, ci sono infiniti punti della successione $z_1, z_2, \cdots, z_n, \cdots$.

Per la dimostrazione denotiamo il rettangolo dato con Δ_1. Lo dividiamo in quattro parti uguali mediante linee parallele agli assi delle coordinate. Almeno una delle parti contiene necessariamente infiniti punti della successione data. Questa parte sarà indicata con Δ_2. Suddividiamo nuovamente il rettangolo Δ_2 in quattro parti uguali e scegliamo tra queste un Δ_3 che contiene infiniti punti della successione data, e così via.

Otteniamo una successione di rettangoli $\Delta_1, \Delta_2, \Delta_3, \cdots$, ciascuno contenuto nel precedente, le cui diagonali tendono a zero. Per la proprietà di continuità esiste un punto z_0 comune a tutti questi rettangoli. È il punto cercato. Infatti, per quanto piccolo sia un intorno di z_0, tutti i rettangoli della successione, a partire da un certo indice, sono contenuti nell'intorno: basta che le loro diagonali siano più piccole del suo raggio. Ciascun rettangolo contiene infiniti punti della successione $z_1, z_2, \cdots, z_n, \cdots$. Il teorema di Bolzano-Weierstrass è così dimostrato.

Ora è facile dimostrare il teorema sul minimo del modulo $|f(z)|$ di un polinomio. Come prima, sia $|f(0)| = g$, sia G un numero maggiore di g, e sia R tale che per $|z| > R$, abbiamo $|f(z)| > G$.

Se $g = 0$, cioè $f(0) = 0$, allora il modulo $|f(z)|$ del polinomio ha minimo nel punto 0, poiché in tutti i punti è $\geqslant 0$.

Se $g > 0$ e $|f(z)| \geqslant g$ per tutti i punti z allora $|f(z)|$ ha ancora un minimo nel punto 0. Sia $g > 0$ ed esistano i punti z, in cui $|f(z)| < g$; poi nella successione dei numeri

$$0, \frac{g}{n}, \frac{2g}{n}, \cdots, \frac{ng}{n} = g \qquad (*)$$

troviamo il massimo $c_n = (i/n)\, g$, tale che tutti i valori $|f(z)| \geqslant c_n$. Per il numero successivo $c'_n = [(i+1)/n]\, g$ della successione (*), esiste almeno un punto z_n tale che $|f(z_n)| < c'_n$.

Facciamo tendere n all'infinito. Per tutti gli n abbiamo $|z_n| \leqslant R$, poiché se $|z_n| > R$, quindi $|f(z_n)|$ sarebbe maggiore di G e di conseguenza maggiore anche di g.

Quindi tutti i punti z_n giacciono all'interno di un rettangolo di lati $2R$ e con centro nell'origine. È possibile che alcuni di questi punti coincidano.

Per il teorema di Bolzano-Weierstrass esiste un punto z_0 tale che ogni suo intorno contiene infiniti punti della successione $z_1, z_2, \cdots, z_n, \cdots$.

Stabiliamo che il punto z_0 fornisce il minimo desiderato di $|f(z)|$

Perché in qualsiasi punto z abbiamo

$$|f(z)| > c_n = c_n' - \frac{g}{n} > |f(z_n)| - \frac{g}{n}$$
$$= |f(z_0)| + [|f(z_n)| - |f(z_0)|] - \frac{g}{n}.$$

Questa disuguaglianza è valida per qualsiasi n. Se prendiamo per n una successione di valori per cui z_n si avvicina indefinitamente a z_0, allora a causa della continuità di $|f(z)|$, la differenza $|f(z_n)| - |f(z_0)|$ diventa arbitrariamente piccola in valore assoluto con g/n.

Di conseguenza, $f(z) | \geqslant |f(z_0)|$, cioè $|f(z)|$ in realtà ha un minimo nel punto z_0.

Lemma di D'Alembert.

In considerazione del fatto che tutte le altezze t dei punti sulla superficie del modulo M sono non negative, è chiaro che qualsiasi radice del polinomio $f(z)$, cioè qualsiasi punto z del piano complesso in cui il polinomio $f(z)$ stesso (e di conseguenza il suo modulo $|f(z)|$ anch'esso) è uguale a zero, corrisponde ad un minimo della superficie del modulo M. Tuttavia, come ha dimostrato d'Alembert, è vero anche il contrario: ad ogni minimo la superficie M si estende fino al piano complesso stesso, e di conseguenza in quel punto c'è una radice del polinomio $f(z)$. In altre parole, in ogni punto in cui la quota t è positiva e non zero, non esiste alcun minimo della superficie M. Ciò segue dal cosiddetto lemma di d'Alembert:

Se α è un dato numero complesso tale che $f(\alpha) \neq 0$, allora è sempre possibile trovare un numero complesso h con modulo arbitrariamente piccolo, tale che $|f(\alpha + h)| < |f(\alpha)|$.

Dimostrazione.

Consideriamo il polinomio

$$f(\alpha + h) = a_0(\alpha + h)^n + a_1(\alpha + h)^{n-1} + \cdots + a_{n-1}(\alpha + h) + a_n$$

nelle due indeterminate α e h e ordiniamolo in potenze ascendenti di h. In questo polinomio ci sarà un termine che non contiene affatto h, vale a dire

$$a_0\alpha^n + a_1\alpha^{n-1} + \cdots + a_{n-1}\alpha + a_n = f(\alpha) \neq 0,$$

poiché si è assunto che $f(\alpha) \neq 0$. Ci sarà anche un termine con h^n, vale a dire a_0h^n, poiché si è assunto che $a_0 \neq 0$. Per quanto riguarda i termini con potenze intermedie di h, alcuni di essi, e in alcuni casi tutti, potrebbero mancare. Sia la più bassa potenza positiva di h che si verifica in questo polinomio m, dove $1 \leqslant m \leqslant n$ cioè questa espressione avrà la forma

$$f(\alpha + h) = f(\alpha) + Ah^m + Bh^{m+1} + Ch^{m+2} + \cdots + a_0 h^n.$$

Scriviamolo come:

$$f(\alpha + h) = f(\alpha) + Ah^m + Ah^m \left(\frac{B}{A} h + \frac{C}{A} h^2 + \cdots + \frac{a_0}{A} h^{n-m}\right),$$

dove $A \neq 0$ e B, C, ecc., possono o meno essere uguali a zero.

Dopo questa preparazione la dimostrazione del lemma di d'Alembert è la seguente. Per h è sufficiente prendere un numero complesso con modulo così piccolo che la lunghezza del vettore Ah^m sia minore della lunghezza del vettore $f(\alpha)$ e con argomento tale che la direzione del vettore Ah^m sia opposta alla direzione del vettore $f(\alpha)$. Quindi il vettore $f(\alpha) + Ah^m$ sarà più corto del vettore $f(\alpha)$. Ma se il modulo di h è considerato sufficientemente piccolo, il modulo dell'espressione

$$\left(\frac{B}{A} h + \frac{C}{A} h^2 + \cdots + \frac{a_0}{A} h^{n-m}\right)$$

può essere reso arbitrariamente piccolo, ad esempio inferiore a uno, e di conseguenza, il vettore

$$\Delta = Ah^m \left(\frac{B}{A} h + \frac{C}{A} h^2 + \cdots + \frac{a_0}{A} h^{n-m}\right)$$

è più corto del vettore Ah^m e quindi anche il vettore $f(\alpha + h) = f(\alpha) + Ah^m + \Delta$, come si vede in (figura 5), è più corto del vettore $f(\alpha)$, anche se la direzione del vettore Δ è nella direzione opposta del vettore Ah^m.

I dettagli di questa prova sono i seguenti:

1. Poiché nella moltiplicazione vengono aggiunti gli argomenti dei fattori, dobbiamo prendere l'argomento di h tale che

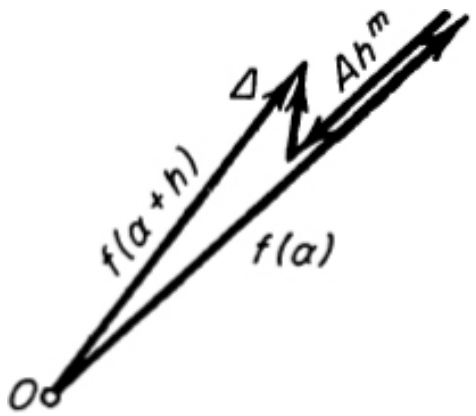


Figura 5.

$$\arg A + m \cdot \arg h = \arg f(\alpha) + 180°,$$

cioè, è necessario prendere

$$\arg h = \frac{\arg f(\alpha) - \arg A + 180^\circ}{m}.$$

2. Il modulo di

$$\left(\frac{B}{A}h + \frac{C}{A}h^2 + \cdots + \frac{a_0}{A}h^{n-m}\right)$$

non è maggiore della somma dei moduli dei suoi addendi

$$T = \left|\frac{B}{A}h\right| + \left|\frac{C}{A}h^2\right| + \cdots + \left|\frac{a_0}{A}h^{n-m}\right|;$$

Inoltre, diminuendo il modulo di h, ciascun addendo di questa somma, e quindi la somma stessa, può essere reso arbitrariamente piccolo. Esiste dunque un numero $h_0 > 0$ tale che, scegliendo h con l'argomento sopra indicato e con $0 < |h| < h_0$, sono soddisfatte entrambe le condizioni $|Ah^m| < |f(\alpha)|$ e $T < 1$. Per tale h avremo $|f(\alpha + h)| < |f(\alpha)|$. Il lemma di d'Alembert è così dimostrato.

Dal lemma di d'Alembert segue immediatamente che ogni minimo della superficie del modulo M del polinomio $f(z)$ dà una radice di questo polinomio. Infatti, se nel punto α, $f(\alpha) \neq 0$, allora in virtù del lemma di d'Alembert ai punti arbitrariamente vicini $\alpha + h$ avremmo $|f(\alpha + h)| < |f(\alpha)|$, cioè non esisterebbe un cerchio con centro in α, in tutti i cui punti il modulo di $f(z)$ non è minore del modulo di $f(\alpha)$, e quindi nel punto α non avremmo un minimo del modulo di $f(z)$. Con ciò si dimostra il teorema fondamentale dell'algebra.

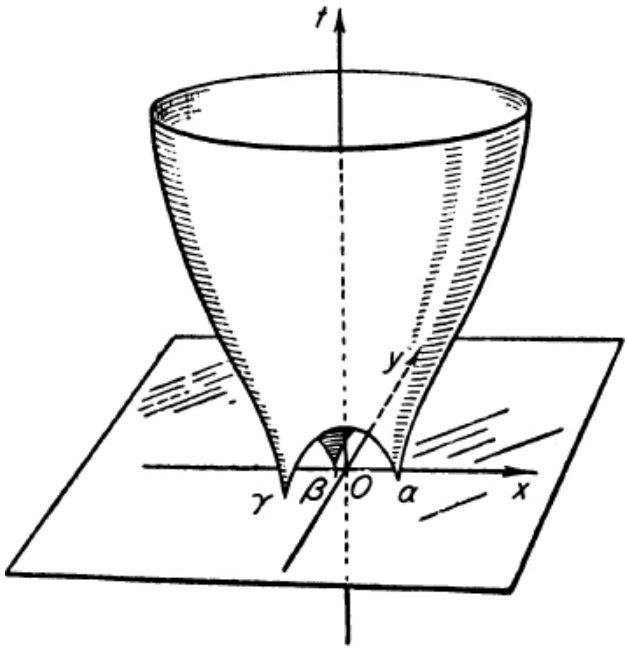


Figura 6.

La forma generale della superficie del modulo M.

La superficie del modulo *M* del polinomio *f*(*z*) si trova sopra il piano complesso z. Ha la forma mostrata in figura 6. Si può dimostrare che a quote maggiori *t*, la superficie *M* differisce molto poco dalla superficie ottenuta ruotando la parabola di *n*° grado $t = | a_0 | x^n$ attorno all'asse Ot. Ma per i piccoli *t* la superficie *M* ha dei minimi, il cui numero è pari al numero di radici distinte dell'equazione *f*(*z*) = 0. A tutti questi minimi la superficie *M* tocca il piano complesso *z* stesso.

§ 4. Studio della distribuzione delle radici di un polinomio nel piano complesso

A questa domanda sono collegati numerosi problemi pratici importanti: senza risolvere una determinata equazione, ottenere alcune informazioni sulla distribuzione delle sue radici sul piano complesso. Il primo problema, storicamente, è stato quello di determinare il numero di radici reali di un'equazione. Cioè, se è data un'equazione a coefficienti reali, allora mediante qualche test dipendente dai suoi coefficienti, per determinare, senza risolvere l'equazione, se ha radici reali e, se le ha, quante; ovvero quante radici positive e quante negative ha; ovvero quante radici reali si trovano tra i limiti dati *a* e *b*.

Derivate di un polinomio.

In questa sezione un ruolo essenziale sarà giocato dalla derivata di un polinomio. La definizione di derivata di una funzione è stata data nel Capitolo II.

Per il polinomio $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}, x + a_n$, la derivata è data, come è noto, dal polinomio

$$na_0x^{n-1} + (n-1)\,a_1x^{n-2} + \cdots + a_{n-1}\,.$$

Il concetto di derivata nel Capitolo II è stato considerato solo per funzioni di variabile reale. In algebra è necessario considerare la variabile come se assumesse valori complessi arbitrari e introdurre polinomi a coefficienti complessi.

Tuttavia, si può mantenere la precedente definizione di derivata, cioè come limite del rapporto tra l'incremento della funzione e l'incremento della variabile indipendente. La formula per calcolare la derivata di un polinomio a coefficienti complessi e le leggi fondamentali per la derivata di somma, prodotto e potenza rimangono le stesse di prima.[9]

Radici semplici e multiple di un polinomio.

Nel §2 di questo Capitolo si è stabilito che se il numero *a* è radice del polinomio *f*(*x*), allora *f*(*x*) è divisibile per *x* - *a* senza resto. Se *f*(*x*) non è divisibile per $(x - a)^2$, il numero *a* è chiamato radice semplice del polinomio *f*(*x*). In generale, se il polinomio *f*(*x*) è divisibile per $(x - a)^k$ ma non per $(x - a)^{k+1}$, allora il numero a è chiamato radice di molteplicità *k*.

Una radice a di molteplicità k è spesso considerata come k radici. La base di ciò è che il fattore $(x - a)^k$, presente nella fattorizzazione di $f(x)$ in fattori lineari, è il prodotto di k fattori, ciascuno uguale a $(x - a)$.

In virtù del fatto che ogni polinomio di grado n è scomponibile nel prodotto di n fattori lineari, il numero di radici del polinomio è pari al suo grado, se si tiene conto della molteplicità di ciascuna radice.

Sono veri i seguenti teoremi:

1. Una radice semplice di un polinomio non è una radice della sua derivata.

2. Una radice multipla di un polinomio è una radice della sua derivata con molteplicità diminuita di uno.

Infatti, sia $f(x) = (x - a)^k f_1(x)$ e supponiamo che $f_1(x)$ non sia divisibile per $(x - a)$, ovvero $f_1(a) \neq 0$. Quindi

$$f'(x) = k(x-a)^{k-1} f_1(x) + (x-a)^k f_1'(x) = (x-a)^{k-1}\,[kf_1(x) + (x-a) f_1'(x)] = (x-a)^{k-1} F(x).$$

Il polinomio $F(x) = kf_1(x) + (x - a) f'_1(x)$ non è divisibile per $(x - a)$, poiché $F(a) = kf_1(a) \neq 0$.

Di conseguenza, $f'(x)$ per $k = 1$ non è divisibile per $x - a$, e per $k > 1 f'(x)$ è divisibile per $(x - a)^{k-1}$ ma non per $(x - a)^k$. Con questo sono dimostrati entrambi i teoremi.

Teorema di Rolle e alcune sue conseguenze.

Secondo il noto teorema di Rolle,[10] se i numeri reali a e b sono radici di un polinomio a coefficienti reali, allora esiste un numero α compreso tra a e b che è radice della derivata.

Dal teorema di Rolle seguono i seguenti interessanti teoremi:

1. Se tutte le radici del polinomio $f(x) = a_0 x^n + \cdots + a_n$ sono reali, allora anche tutte le radici della sua derivata sono reali. Inoltre, tra due radici adiacenti di $f(x)$ esiste una radice di $f'(x)$ e questa radice è semplice. Infatti, siano $x_1 < x_2 < \cdots < x_k$, le radici di $f(x)$ con molteplicità m_1, m_2, $\cdots$, m_k, rispettivamente. Chiaramente, $m_1 + m_2 + \cdots + m_k = n$.

Allora la derivata $f'(x)$, per il teorema sopra su radici multiple, avrà radici x_1, x_2, $\cdots$, x_k con molteplicità $m_1 - 1$, $m_2 - 1$, $\cdots$, $m_k - 1$, e per il teorema di Rolle esiste almeno una radice y_1, y_2, $\cdots$, y_{k-1} all'interno di ciascuno degli intervalli (x_1, x_2), (x_2, x_3), $\cdots$, (x_{k-1}, x_k) tra due radici successive di $f(x)$. Pertanto, il numero di radici reali di $f'(x)$ è uguale (per quanto riguarda le molteplicità) almeno a $(m_1 - 1) + (m_2 - 1) + \cdots + (m_k - 1) + k - 1 = n - 1$. Ma $f'(x)$ come polinomio di grado $n - 1$ ha (rispetto alle molteplicità) $n - 1$ radici. Di conseguenza, tutte le radici di $f'(x)$ sono reali, y_1, y_2, $\cdots$, y_{k-1} sono radici semplici, e le radici diverse da x_1, x_2, $\cdots$, x_k e y_1, y_2, $\cdots$, y_{k-1} del polinomio $f'(x)$ non esistono.

2. Se tutte le radici di un polinomio $f(x)$ sono reali e di queste p sono positive, allora $f'(x)$ ha p o $p - 1$ radici positive.

Infatti, siano $x_1 < x_2 < \cdots < x_k$ tutte radici positive del polinomio $f(x)$ con molteplicità m_1, m_2, $\cdots$, m_k, rispettivamente. Quindi $m_1 + m_2 + \cdots + m_k = p$. La derivata $f'(x)$ avrà le seguenti radici positive: x_1, x_2, $\cdots$, x_k con molteplicità $m_1 - 1$, $m_2 - 1$, $\cdots$, $m_k - 1$; radici semplici y_1, y_2, $\cdots$, y_{k-1}, giacenti negli intervalli (x_1, x_2), $\cdots$, (x_{k-1}, x_k); e può anche avere una radice semplice y_0 situata nell'intervallo (x_0, x_1) dove x_0 è la più grande radice non positiva di $f(x)$. Di conseguenza, il numero di radici positive è pari a $(m_1 - 1) + \cdots + (m_k - 1) + k - 1 = p - 1$ oppure $(m_1 - 1) + \cdots + (m_k - 1) + (k - 1) + 1 = p$ che doveva essere dimostrato.

Legge dei segni di Cartesio.

Nel suo importante libro del 1637 “Geometria”, in cui fu data la prima presentazione della geometria analitica, Cartesio, tra le altre cose, fornì il primo teorema algebrico significativo riguardante la distribuzione delle radici di un polinomio sul piano complesso, la cosiddetta “legge dei segni di Cartesio”. Si può affermare come segue:

Se i coefficienti di un'equazione sono reali e si sa che anche tutte le sue radici sono reali, allora il numero delle sue radici positive, tenendo conto delle molteplicità, è uguale al numero di cambiamenti di segno nella successione dei suoi coefficienti. Se ha anche radici complesse, allora questo numero è uguale o inferiore di un numero pari al numero di questi cambiamenti di segno.

Spieghiamo innanzitutto cosa intendiamo per numero di cambiamenti di segno nella successione dei coefficienti dell'equazione. Per ottenere questo numero scriviamo tutti i coefficienti dell'equazione, ad esempio in ordine di potenze decrescenti dell'incognita, compreso il coefficiente di x^n e il termine costante, ma omettendo i coefficienti pari a zero, e consideriamo tutte le coppie di numeri successivi della successione così ottenuta. Se in una coppia del genere i segni dei numeri sono diversi, allora lo chiamiamo cambio di segno. Ad esempio, se l'equazione data è

$$x^7 + 3x^5 - 5x^4 - 8x^2 + 7x + 2 = 0$$

quindi la successione dei suoi coefficienti è

$$1, 3, -5, -8, 7, 2$$

e ci sono 2 cambi di segno.

Passiamo ora alla dimostrazione della prima parte del teorema.[11]

Senza perdita di generalità possiamo assumere che il coefficiente direttivo a_0 del polinomio $f(x) = a_0x^n + \cdots + a_n$ sia positivo.

Innanzitutto stabiliamo che se, $f(x)$ ha solo radici reali e di queste p sono positive (contando le molteplicità) allora $(-1)^p$ è il segno dell'ultimo coefficiente di $f(x)$ diverso da zero.

Infatti, sia

$$f(x) = a_0x^n + \cdots + a_{n-k}x^k$$
$$= a_0x^k(x - x_1) \cdots (x - x_p)(x - x_{p+1}) \cdots (x - x_{n-k}).$$

dove $x_1, \cdots, x_p$ sono le radici positive e $x_{p+1}, \cdots, x_{n-k}$ quelle negative, contate con le rispettive molteplicità. Allora $a_{n-k} = a_0(-1)^p x_1 \cdots x_p(-x_{p+1}) \cdots (-x_{n-k})$. Poiché tutti i fattori, salvo $(-1)^p$, sono positivi, il segno di a_{n-k} è $(-1)^p$.[12]

La dimostrazione successiva si basa sul metodo dell'induzione matematica.

Per i polinomi di primo grado il teorema è banale. Infatti, un polinomio di primo grado $a_0x + a_1$, ha un'unica radice $-a_1/a_0$ che è positiva se e solo se a_0 e a_1 hanno segni opposti.

Supponiamo ora che il teorema sia dimostrato per tutti i polinomi di grado $n - 1$ con radici reali, e con questa assunzione lo dimostreremo per ogni polinomio $f(x) = a_0x^n + \cdots + a_{n-1}x + a_n$, di grado n.

1. $a_n = 0$. Consideriamo il polinomio $f_1(x) = a_0x^{n-1} + \ldots + a_{n-1}$. Le radici positive dei polinomi $f(x)$ e $f_1(x)$ sono le stesse; anche il numero di cambiamenti di segno nella successione dei loro coefficienti è lo stesso. Per il polinomio $f_1(x)$ vale la legge di Cartesio; di conseguenza vale per il polinomio $f(x)$.

2. $a_n \neq 0$. Consideriamo la derivata

$$f'(x) = na_0x^{n-1} + (n-1)a_1x^{n-2} + \cdots + a_{n-1}.$$

È chiaro che il numero di cambiamenti di segno nella successione dei coefficienti del polinomio $f(x)$ è uguale al numero analogo per la derivata $f'(x)$, se i segni di a_n e l'ultimo coefficiente diverso da zero della derivata coincidono, oppure è uno in più, se i segni sono opposti.

Per quanto detto sopra all'inizio della dimostrazione, nel primo caso il numero di radici positive di $f(x)$ e di $f'(x)$ hanno la stessa parità (sono entrambi pari o entrambi dispari), e nel secondo caso hanno parità opposta. Ma come abbiamo dedotto dal teorema di Rolle, il numero di radici positive di un polinomio, se tutte le sue radici sono reali, può essere uguale al numero di radici positive della sua derivata, oppure essere uno di più. Tenendo presente ciò, notiamo che nel primo caso $f(x)$ ha lo stesso numero di radici positive di $f'(x)$, e nel secondo caso una in più. Per $f'(x)$ vale la legge di Cartesio mediante l'ipotesi di induzione, cioè il numero di radici positive di $f'(x)$ è uguale al numero di cambiamenti di segno nella successione dei suoi coefficienti. Di conseguenza, in entrambi i casi il numero di radici positive di $f(x)$ è pari al numero di cambiamenti di segno nella successione dei coefficienti, e questa è la prova richiesta.

La seconda parte della legge di Cartesio non è più complicata da stabilire, e qui ne ometteremo la dimostrazione:

Osservazione 1.

La prima affermazione del teorema di Cartesio è particolarmente importante, poiché in molti problemi pratici è automaticamente noto se tutte le radici di una data equazione sono reali. In questo caso si può determinare rapidamente quante radici sono positive e quante negative. Inoltre si può vedere subito quante radici zero ha l'equazione.

Osservazione 2.

Se nel polinomio dato poniamo $x = y + a$ dove a è un numero reale arbitrario dato, cioè formiamo il polinomio $f(y + a)$, allora le radici positive y di questo polinomio saranno quelle e solo quelle ottenuti dalle radici x del polinomio dato $f(x)$ maggiori di a. Pertanto il numero di radici del polinomio dato $f(x)$, le cui radici sono tutte reali, comprese tra i limiti dati a e b ($b > a$), è uguale al numero di cambiamenti di segno del polinomio $f(y + a)$ meno il numero di cambi di segno per il polinomio $f(z + b)$. Se, tuttavia, non tutte le radici di $f(x)$ sono reali, allora si può dimostrare che questo numero è uguale a questa differenza o inferiore a essa di un numero pari. Questo è il cosiddetto teorema di Budan.

Il teorema di Sturm.

La legge dei segni di Cartesio, così come il teorema di Budan, tuttavia, non danno una risposta al problema: una data equazione a coefficienti reali ha almeno una radice reale, quante radici reali ha in tutto e quante radici reali ha tra i limiti dati a e b? Per più di due secoli i matematici tentarono di risolvere questi problemi ma senza risultato. Una lunga serie di sforzi in questa direzione furono fatti da Cartesio, Newton, Sylvester, Fourier e molti altri, ma non riuscirono a risolvere nemmeno il primo di questi problemi, finché, finalmente, nel 1835, il matematico francese Sturm suggerì un metodo che risolvesse tutti e tre i problemi.

In realtà il metodo di Sturm non è molto complicato, ma è di tale natura che si potrebbe cercarlo a lungo e non trovarlo. Lo stesso Sturm era molto felice di essere riuscito a risolvere questo notevole ed estremamente importante problema pratico di algebra. Nelle sue lezioni, quando arrivava alla presentazione del suo risultato, era solito dire: “Ecco il teorema di cui porto il nome!”. Ma va detto che Sturm non risolse questo problema per puro caso; rifletté per molti anni su questioni ad esso relative.

Sia $f(z)$ un polinomio a coefficienti reali e $f_1(z)$ sia la derivata $f'(z)$. Dividiamo il polinomio $f(z)$ per $f_1(z)$ e denotiamo il resto di questa divisione con $f_2(z)$, prendendolo con il segno opposto. Quindi dividiamo $f_1(z)$ per $f_2(z)$ e denotiamo il resto, preso con segno opposto, con $f_3(z)$, ecc.

Se il polinomio non ha radici multiple, si può dimostrare che l'ultimo polinomio diverso da zero $f_l(z)$ della successione costruita sarà un numero costante c.

Il teorema di Sturm è il seguente: Se $a < b$ sono due numeri reali, che non sono radici del polinomio $f(z)$, allora sostituendo nei polinomi

$$f(z), f_1(z), \cdots, f_{s-1}(z), c$$

$z = a$ e $z = b$, otteniamo due successioni di numeri reali tali che il numero di cambi di segno nella successione (I) è maggiore o uguale al numero di cambi di segno nella successione (II) e la differenza tra questi numeri di cambi di segno è esattamente uguale al numero di radici reali di $f(z)$ comprese tra a e b, ovvero il numero di queste radici è pari alla perdita di cambi di segno nella successione (I) nel passare da a a b.

$$f(a), f_1(a), f_2(a), \cdots, f_{s-1}(a), c, \qquad \text{(I)}$$

$$f(b), f_1(b), f_2(b), \cdots, f_{s-1}(b), c, \qquad \text{(II)}$$

La dimostrazione del teorema di Sturm non è più difficile della dimostrazione del teorema di Cartesio, ma non la daremo qui.

Il teorema di Sturm permette di calcolare il numero di radici di un polinomio a coefficienti reali su qualsiasi segmento dell'asse reale. Pertanto l'applicazione del teorema di Sturm a qualsiasi polinomio dato fornisce un quadro chiaro della distribuzione delle radici di un polinomio sull'asse reale, in particolare ci permette di separare le radici, cioè di costruire segmenti in ciascuno dei quali è contenuta solo una radice del polinomio.

In molte applicazioni è altrettanto importante la soluzione dell'analogo problema per le radici complesse di un polinomio. Poiché i numeri complessi sono rappresentati da punti non sulla retta ma sul piano, è impossibile parlare di "segmenti" in cui sono contenute radici complesse; invece che un segmento dobbiamo considerare una regione, cioè una parte del piano, scelta in un modo o nell'altro.

Pertanto, rispetto alle radici complesse si pone il seguente problema:

Dato un polinomio $f(z)$ e una regione nel piano complesso, è necessario trovare il numero di radici del polinomio all'interno di tale regione.

Assumiamo che la regione sia delimitata da un contorno chiuso (figura 7) e che sul contorno il polinomio $f(z)$ non abbia radici.

Immaginiamo che il punto z percorra il contorno della regione una volta in direzione positiva. Ogni valore del polinomio è rappresentato anche da punti sul piano. Con il cambiamento continuo di z anche il polinomio $f(z)$ cambia continuamente. Pertanto, mentre z percorre il contorno della regione, $f(z)$ descrive una curva chiusa.

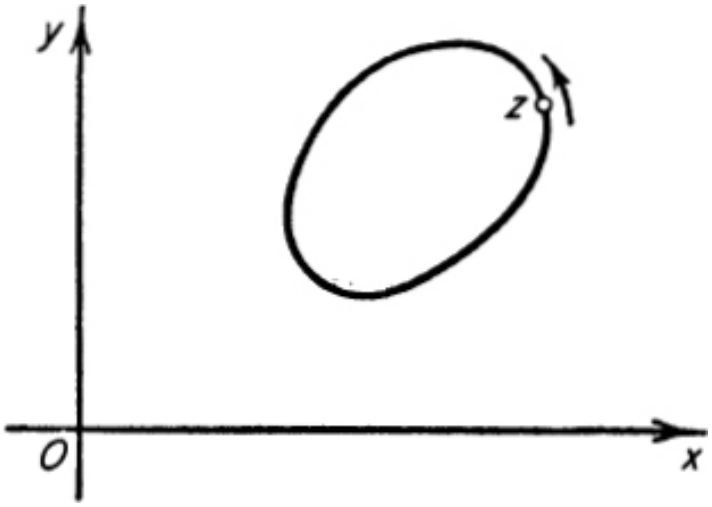


Figura 7.

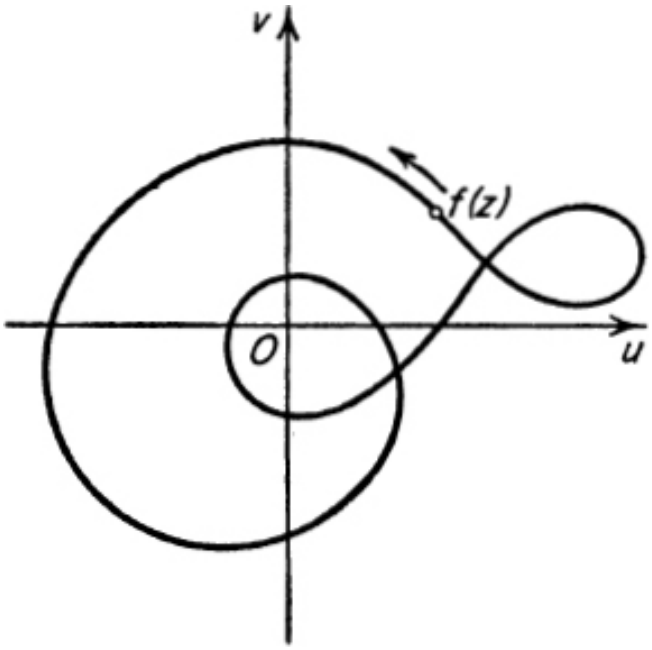


Figura 8.

Questa curva non passerà per l'origine del sistema di coordinate, poiché $f(z)$ per ipotesi non si riduce a zero in nessuno dei punti del contorno (figura 8).

La risposta al problema sopra menzionato è data dal seguente teorema:

Principio dell'argomento.

Il numero di radici del polinomio $f(z)$ all'interno della regione delimitata da una curva chiusa C è uguale al numero di volte che il punto $f(z)$ avvolge l'origine mentre z gira attorno al contorno C una volta nella direzione positiva.

Per la dimostrazione scomponiamo $f(z)$in fattori lineari

$$f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n = a_0(z - z_1)(z - z_2) \cdots (z - z_n).$$

Sappiamo che l'argomento del prodotto di più numeri complessi è uguale alla somma degli argomenti dei fattori. Di conseguenza,

$$\arg f(z) = \arg a_0 + \arg (z - z_1) + \arg (z - z_2) + \cdots + \arg (z - z_n).$$

Indichiamo con $\Delta \arg f(z)$l'incremento dell'argomento di $f(z)$, calcolato assumendo che z giri una volta attorno al contorno C. È chiaro che $\Delta \arg f(z)$è 2π moltiplicato per il numero di volte in cui il punto $f(z)$avvolge attorno all'origine.

Chiaramente,

$$\begin{aligned}\Delta \arg f(z) = {} & \Delta \arg a_0 + \Delta \arg (z - z_1) \\ & + \Delta \arg (z - z_2) + \cdots + \Delta \arg (z - z_n).\end{aligned}$$

È chiaro che Δ arg a_0 = 0 poiché a_0 è una costante. Quindi, $z - z_1$ è rappresentato dal vettore che va dal punto z_1, al punto z. Supponiamo che z_1 si trovi all'interno della regione. Geometricamente è chiaro (figura 9) che man mano che il punto z aggira il contorno C il vettore $z - z_1$ compie un giro completo attorno al suo punto iniziale, per cui Δ arg $(z - z_1) = 2\pi$. Assumiamo ora che il punto z_2 si trovi all'esterno della regione. In questo caso il vettore "oscilla" da un lato e dall'altro, e ritorna nella sua posizione originale senza fare una rivoluzione attorno al suo punto iniziale, per cui Δ arg $(z - z_2) = 0$. Possiamo ragionare allo stesso modo su tutte le radici. Di conseguenza Δ arg $f(z)$ è uguale a 2π moltiplicato per il numero di radici di $f(z)$ che si trovano all'interno della regione. Quindi il numero di radici di $f(z)$ all'interno della regione è uguale al numero di volte che il punto $f(z)$ si avvolge attorno all'origine, e questa è la prova richiesta.

Questo teorema ci permette di risolvere il problema in ogni caso particolare e di disegnare la curva tracciata dal punto $f(z)$ con qualsiasi grado di precisione. Per fare ciò è necessario prendere un insieme sufficientemente denso di punti z sul contorno C, calcolare i corrispondenti valori $f(z)$ e unirli tramite una curva continua. Tuttavia, in alcuni casi possiamo farcela senza questi noiosi calcoli. Indichiamo uno dei metodi con un esempio numerico.

Esempio.

Troviamo il numero di radici del polinomio $f(z) = z^{11} + 5z^2 - 2$ all'interno di una circonferenza di raggio 1 con centro nell'origine.

Sul cerchio indicato $| z | = 1$, uno dei tre termini che compongono il polinomio $f(z)$, cioè $5z^2$, domina sugli altri. Infatti, $| 5z^2 | = 5$, ma $| z^{11} - 2 | \leqslant | z |^{11} + 2 = 3$. Questa proprietà ci permette di ragionare così. Indichiamo $z^{11} + 5z^2 - 2$ con w, $5z^2$ con N_1 e $z^{11} - 2$ con N_2. Mentre il punto z compie un giro attorno alla circonferenza unitaria, $N_1 = 5z^2$ avvolge due volte attorno ad una circonferenza di raggio 5, poiché $| N_1 | = 5$ e arg $N_1 = 2$ arg z. Il punto w è "legato" al punto N_1 da un vettore la cui lunghezza è $| N_2 |$ | $N_2 \leqslant 3$, cioè la distanza dal punto w al punto N_1 è sempre inferiore alla distanza da N_1 all'origine del sistema di coordinate.

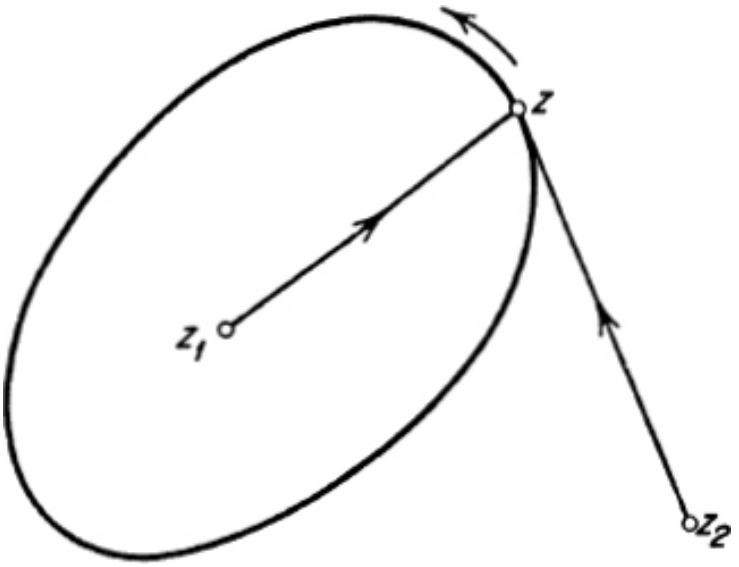


Figura 9.

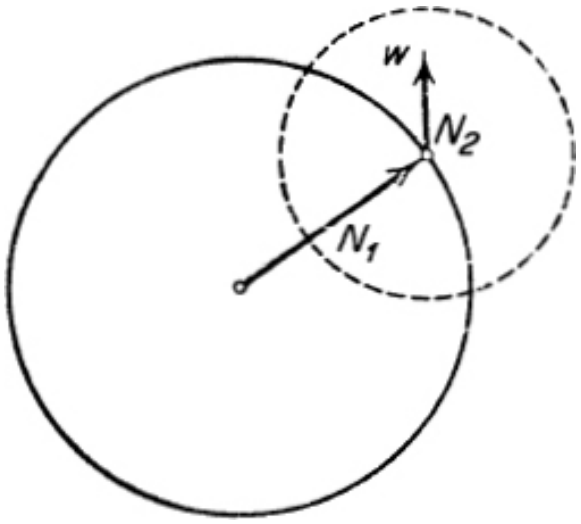


Figura 10.

Di conseguenza il punto w, per quanto "avvolga" attorno a N_1 (figura 10), non può "indipendentemente" aggirare l'origine, e quindi avvolge attorno all'origine esattamente tante volte quanto fa il punto N_1, cioè due volte. Di conseguenza, il numero di radici di $f(z)$ all'interno della regione in questione è pari a due.

Il problema di Hurwitz.

In meccanica, in particolare nella teoria delle oscillazioni e del controllo, un ruolo importante è giocato dalle condizioni che permettono di decidere se tutte le radici di un dato polinomio $f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n$, (con coefficienti reali) hanno parti reali negative, cioè giacciono nel semipiano a sinistra dell'asse immaginario.

Uno dei criteri per risolvere questo problema è facile da ottenere da ragioni simili al principio dell'argomento.

Assumeremo che $a_0 > 0$.

Lasciamo che il punto z (figura 11) si sposti sull'asse immaginario dall'alto verso il basso, ovvero sia $z = iy$ poiché y cambia da $+\infty$ a $-\infty$, rimanendo reale. Quindi $f(z)$ descrive una curva con rami infiniti. Per la nostra indagine la curva strettamente correlata descritta dalla funzione

$$f_1(z) = (i)^{-n} f(z) = a_0 y^n - a_2 y^{n-2} + a_4 y^{n-4} + \cdots - i(a_1 y^{n-1} - a_3 y^{n-3} + \cdots)$$
$$= \phi(y) - i\psi(y),$$

Dove

$$\phi(y) = a_0 y^n - a_2 y^{n-2} + \cdots,$$
$$\psi(y) = a_1 y^{n-1} - a_3 y^{n-3} + \cdots$$

è più conveniente.

Poiché $\arg i = \pi/2$, quindi $\arg f_1(z) = -n\pi/2 + \arg f(z)$, e di conseguenza gli incrementi degli argomenti di $f(z)$ e $f_1(z)$ sono uguali. Calcoliamo l'incremento dell'argomento del punto $f_1(z)$ mentre z si sposta sull'asse immaginario verso il basso.

Sia $f(z) = a_0(z - z_1)(z - z_2) \cdots (z - z_n)$. Poi

$$\arg f_1(z) = \arg(a_0 i^{-n}) + \arg(z - z_1) + \arg(z - z_2) + \cdots + \arg(z - z_n).$$

È geometricamente chiaro che l'incremento di $\arg(z - z_k)$ è pari a π, se z_k giace nel semipiano destro e a $-\pi$ se z_k si trova nel semipiano sinistro (figura 11).

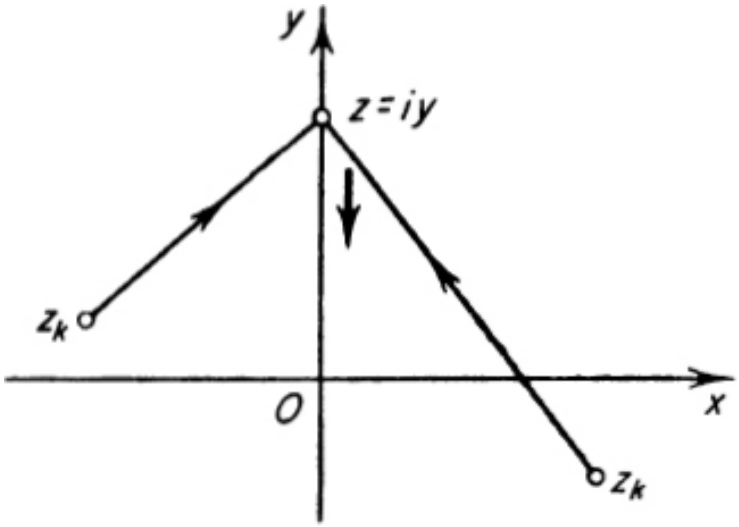


Figura 11.

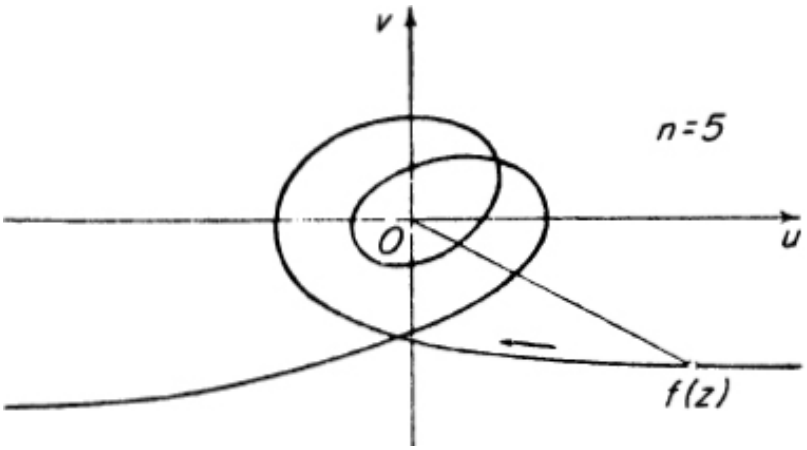


Figura 12.

Pertanto l'incremento dell'argomento di $f_1(z)$ è uguale a $\pi(N_1 - N_2)$, dove N_1 è il numero di radici di, $f(z)$ nel semipiano destro e N_2 è il numero di radici nel semipiano sinistro. Affinché tutte le radici giacciano nel semipiano sinistro è necessario e sufficiente che l'incremento dell'argomento del punto $f_1(z)$ sia uguale a $-\pi n$, cioè che il punto $f_1(z)$ faccia n mezzi giri in senso orario attorno all'origine (figura 12).

Notiamo che il punto $f_1(z) = \phi(y) - i\psi(y)$ interseca l'asse immaginario per quei valori di y che sono radici di $\phi(y)$ e l'asse reale per le radici di $\psi(y)$. Poiché $\phi(y)$ non ha più di n radici reali e il numero di radici reali di $\psi(y)$ non è maggiore di $n - 1$, è facile vedere geometricamente che $f_1(z)$ può compiere n mezzi giri in senso orario se e solo se la curva proviene dal quarto quadrante e interseca a sua volta la metà negativa dell'asse immaginario, la metà negativa dell'asse reale, la metà positiva dell'asse immaginario, la metà positiva dell'asse reale, ecc., in modo che il numero totale di punti di intersezione con l'asse immaginario è pari a n (uno per ogni mezzo giro), e con l'asse reale è pari a $n - 1$ (uno in meno del numero di mezzi giri). Pertanto il coefficiente a_1 deve essere positivo, e le radici dei polinomi $\phi(y)$ e $\psi(y)$ devono essere tutte reali ed alternate. Quest'ultima affermazione significa che se $y_1 > y_2 > \cdots > y_n$ sono le radici di $\phi(y)$ disposte in ordine decrescente, e $\eta_1 > \eta_2 > \cdots > \eta_{n-1}$ sono le radici di $\psi(y)$, quindi $y_1 > \eta_1 > y_2 > \eta_2 > \cdots > y_{n-1} > \eta_{n-1} > y_n$.

Pertanto, affinché tutte le radici del polinomio $f(z) = a_0 z^n + a_1 z^{n-1} + \cdots + a_n$ con coefficienti reali e $a_0 > 0$ giacciano nel semipiano sinistro, è necessario e sufficiente che il coefficiente a_1 sia positivo e le radici dei polinomi $\phi(y) = a_0 y^n - a_2 y^{n-2} + a_4 y^{n-4} - \cdots$ e $\psi(y) = a_1 y^{n-1} - a_3 y^{n-3} + \cdots$ siano tutte reali e alternate.

Questa condizione equivale alla ben nota condizione di Hurwitz secondo la quale tutti i seguenti determinanti sono positivi:

$$a_1\,,\quad \begin{vmatrix} a_1 & a_0 \\ a_3 & a_2 \end{vmatrix},\quad \begin{vmatrix} a_1 & a_0 & a_{-1} \\ a_3 & a_2 & a_1 \\ a_5 & a_4 & a_3 \end{vmatrix},\ \cdots,\quad \begin{vmatrix} a_1 & a_0 & a_{-1} & \cdots & a_{2-n} \\ a_3 & a_2 & a_1 & \cdots & a_{4-n} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{2n-1} & a_{2n-2} & a_{2n-3} & \cdots & a_n \end{vmatrix},$$

dove tutti i a_i con indici minori di 0 o maggiori di n sono sostituiti da zero (sui determinanti, vedi Capitolo XVI, §3).

§ 5. Calcolo approssimato delle radici

Il metodo di Sturm in combinazione con il limite inferiore della differenza di due radici reali distinte permette di costruire la “separazione” delle radici reali di un polinomio a coefficienti reali, cioè permette di determinare per ciascuna radice i limiti a e b tra i quali si può trovare solo questa radice. Resta da scoprire un metodo adatto per trovare, nel segmento $a < b$, i numeri $\alpha_1 < \alpha_2 < \alpha_3 < \cdots$ e $\beta_1 > \beta_2 > \beta_3 > \cdots$ che convergono il più rapidamente possibile alla radice desiderata, essendo la prima successione un'approssimazione per difetto e la seconda per eccesso. Ciascuna delle due approssimazioni α_k e β_k, differisce chiaramente dalla radice desiderata x di meno della loro differenza $\beta_k - \alpha_k$, poiché la radice si trova tra di loro. Quindi possiamo trovare i limiti superiori dell'errore quando ci fermiamo a una data approssimazione.

Grafico di un polinomio.

Sia il dato polinomio di grado n a coefficienti reali

$$f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n\,.$$

Consideriamo la curva che rappresenta in coordinate rettangolari l'equazione $y = f(x)$, cioè il grafico di questo polinomio. Questa curva è talvolta chiamata parabola di ordine n. Innanzitutto è chiaro che per ogni x reale esiste un unico valore determinato $y = f(x)$; di conseguenza, il grafico f si estende indefinitamente a destra e a sinistra. Inoltre, in caso di cambio continuo di x, $f(x)$ e $f'(x)$ cambiano continuamente, cioè senza salti. Pertanto, il grafico è una curva liscia. Per x grande in valore assoluto il primo termine a_0x^n supera in valore assoluto la somma di tutti i restanti termini, poiché sono tutti di grado inferiore. Da ciò consegue che se n è pari e $a_0 > 0$, allora il grafico f a destra e a sinistra va all'infinito verso l'alto (e se $a_0 < 0$, verso il basso); ma se n è dispari e $a_0 > 0$, allora a destra va verso l'alto e a sinistra verso il basso (se $a_0 < 0$, allora viceversa).

I punti di intersezione del grafico f con l'asse Ox, cioè quei punti in cui $y = f(x) = 0$, corrispondono alle radici reali dell'equazione $f(x) = 0$; non ce ne sono più di n. Ai massimi e minimi del grafico $y = f(x)$, la derivata $f'(x) = 0$; di conseguenza, il numero di massimi e minimi non è maggiore di $n - 1$. Se su qualche intervallo $f''(x) > 0$, lì la derivata prima

aumenta, cioè il grafico è concavo verso l'alto; se $f''(x) < 0$, il grafico è concavo verso il basso. Poiché alcune radici di $f'(x) = 0$ possono essere complesse, il numero di massimi e minimi del grafico f potrebbe essere inferiore a $n - 1$.

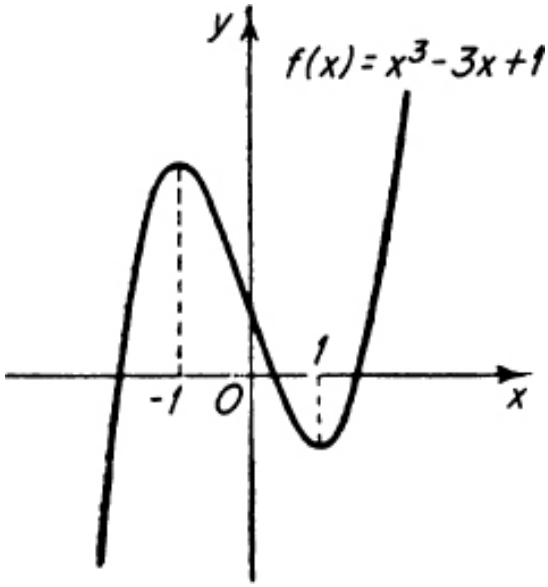


Figura 13.

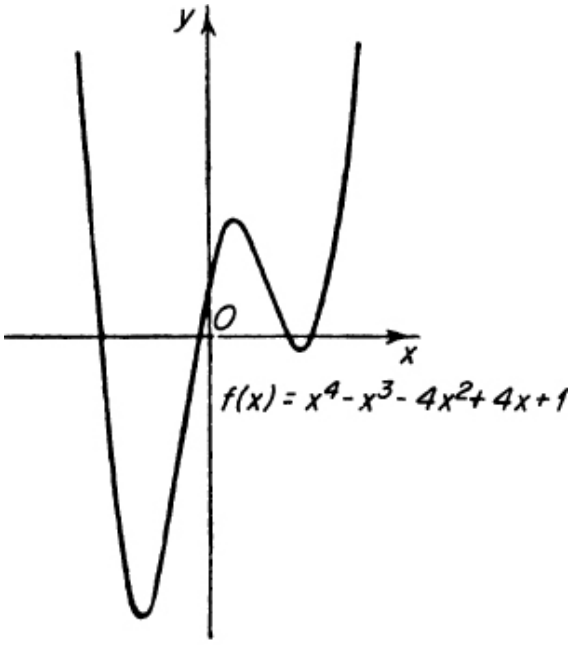


Figura 14.

Ecco alcuni esempi dei grafici dei polinomi

$$f(x) = x^3 - 3x + 1 \quad \text{(figure 13)},$$
$$f(x) = x^4 - x^3 - 4x^2 + 4x + 1 \quad \text{(figure 14)}.$$

Dopo aver costruito il grafico di un polinomio è facile trovare approssimazioni alle sue radici. Vale a dire, le radici sono le ascisse dei punti di intersezione del grafico con l'asse Ox.

Il metodo delle approssimazioni per difetto e per eccesso.

Sostituiamo in $f(x)$ un numero intero, ad esempio 3, e poi 4, 5, ···. Se per 4, 5 e 6 otteniamo lo stesso segno che per 3, mentre per 7 otteniamo il segno opposto, il polinomio ha almeno una radice tra 6 e 7. Sostituiamo ora 6; 6,1; 6,2; ··· e cerchiamo due valori consecutivi, ad esempio 6,4 e 6,5, che diano segni opposti: tra essi vi sarà almeno una radice. Sostituendo poi 6,4; 6,41; 6,42; 6,43; ···, troviamo limiti ancora più vicini, ad esempio 6,42 e 6,43, e così via. Questo è il metodo delle approssimazioni per difetto e per eccesso. Si può semplificarlo notevolmente applicando a ogni passo una trasformazione supplementare del polinomio: dopo il primo passo basterà sostituire numeri interi, anziché frazioni, e precisamente soltanto gli interi 1, 2, ···, 9. Non ci soffermiamo su questa semplificazione.

Il metodo delle tangenti e il metodo delle corde.

Il metodo delle tangenti, detto metodo di Newton, e il metodo delle corde, o dell'interpolazione lineare, detto anche metodo delle false posizioni (regula falsi), vengono utilizzati separatamente o insieme per ottenere stime dell'errore. Supponiamo che tra a e b abbiamo una sola radice del polinomio $f(x)$, per cui $f(a)$ e $f(b)$ siano di segno opposto, e supponiamo anche che la derivata seconda $f''(x)$ tra a e b abbia segno costante. In questo caso la parte del grafico di $f(x)$ compresa tra a e b ha una delle quattro forme (figura 15).

Nei casi I e II della figura 15, la tangente al grafico nel punto di ascissa a interseca l'asse Ox in un punto di ascissa α_1, compresa tra a e la radice cercata. Calcolata α_1, consideriamo la tangente al grafico nel punto con questa ascissa: troviamo così un'ascissa α_2, compresa tra α_1 e la radice cercata. Procedendo nello stesso modo otteniamo α_3, e così via: approssimazioni per difetto sempre migliori. Come mostra la figura, questi valori si avvicinano molto rapidamente alla radice.

Nei casi III e IV occorre invece partire dall'ascissa b, per poi ottenere i punti $\beta_1, \beta_2, \beta_3, \cdots$, cioè approssimazioni sempre migliori per eccesso. Quale dei quattro casi si verifica effettivamente, è facile da determinare dai segni di $f(a)$, $f(b)$ e $f''(x)$ per $a < x < b$. Poiché l'equazione della tangente alla curva $y = f(x)$nei suoi punti con ascissa a è

$$y - f(a) = f'(a)\,(x - a),$$

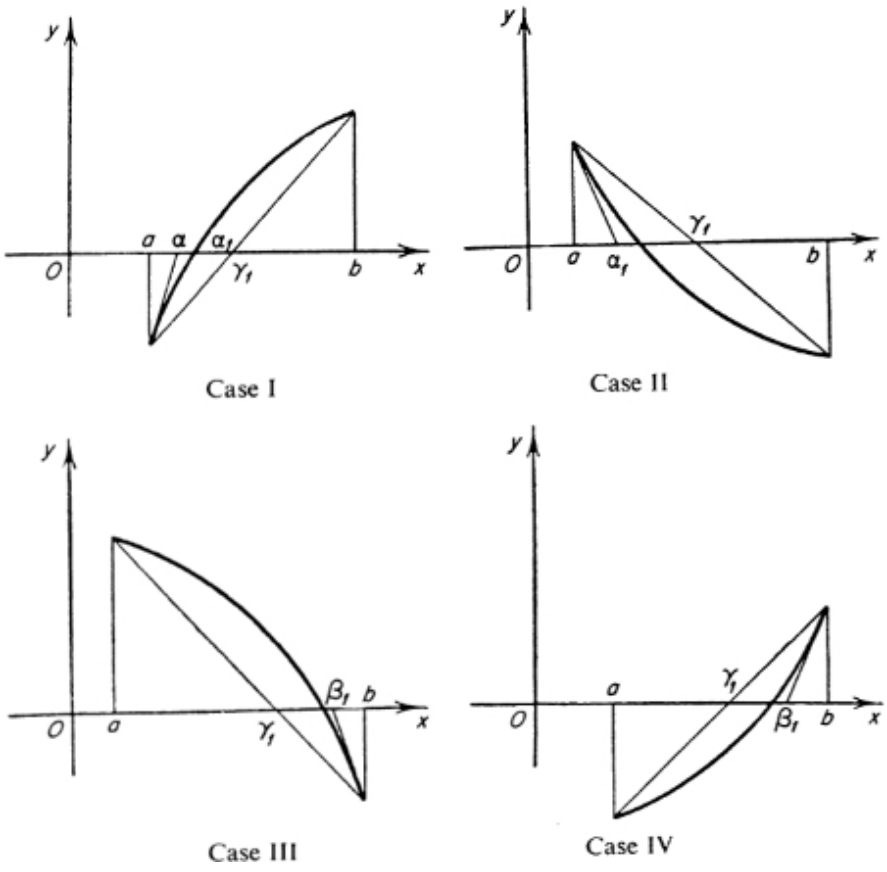


Figura 15. «Case I-IV»: casi I-IV.

dall'uguaglianza si ottiene l'ascissa α_1 del punto della sua intersezione con l'asse Ox

$$0 - f(a) = f'(a)\,(\alpha_1 - a)$$

cioè

$$\alpha_1 = a - \frac{f(a)}{f'(a)}\,.$$

Poi

$$\alpha_2 = \alpha_1 - \frac{f(\alpha_1)}{f'(\alpha_1)}\,, \quad \alpha_3 = \alpha_2 - \frac{f(\alpha_2)}{f'(\alpha_2)}$$

e così via.

Analogamente,

$$\beta_1 = b - \frac{f(b)}{f'(b)}\,, \beta_2 = \beta_1 - \frac{f(\beta_1)}{f'(\beta_1)}\,, \beta_3 = \beta_2 - \frac{f(\beta_2)}{f'(\beta_2)}$$

e così via.

Questo è il metodo di Newton. [13]

Il metodo di interpolazione lineare o falsa posizione consiste in quanto segue. L'equazione di una corda, come l'equazione di una retta passante per due punti dati, ha la forma

$$\frac{x-a}{b-a} = \frac{y-f(a)}{f(b)-f(a)},$$

e l'ascissa γ_1 del punto della sua intersezione con l'asse Ox, come ottenuto dall'equazione

$$\frac{x-a}{b-a} = \frac{0-f(a)}{f(b)-f(a)}$$

è uguale a

$$\gamma_1 = -\frac{(b-a)f(a)}{f(b)-f(a)} + a = \frac{af(b)-bf(a)}{f(b)-f(a)}.$$

Prendendo questo numero per il nuovo b nei casi I e II e per il nuovo valore di a nei casi III e IV, troviamo nei casi I e II

$$\gamma_2 = \frac{af(\gamma_1)-\gamma_1 f(a)}{f(\gamma_1)-f(a)}, \quad \gamma_3 = \frac{af(\gamma_2)-\gamma_2 f(a)}{f(\gamma_2)-f(a)}$$

e così via.

Nei casi III e IV, prendendo γ_1 per il nuovo valore di a troviamo

$$\gamma_2 = \frac{\gamma_1 f(b)-bf(\gamma_1)}{f(b)-f(\gamma_1)}, \quad \gamma_3 = \frac{\gamma_2 f(b)-bf(\gamma_2)}{f(b)-f(\gamma_2)}$$

e così via.

La combinazione di questi due metodi è particolarmente importante poiché (come si vede dai diagrammi) permette, se si conoscono le approssimazioni dall'alto e dal basso, di stimare l'errore, che chiaramente non è maggiore della differenza tra queste approssimazioni, poiché la radice desiderata è tra di loro.

Osservazione.

È importante notare che il fatto che $f(x)$ sia un polinomio, e non qualche altra funzione di x, non ha alcun ruolo né nel metodo di Newton, né nel metodo dell'interpolazione lineare, cioè entrambi questi metodi e la loro combinazione possono essere adattati, nelle condizioni sopra menzionate, alle equazioni trascendenti.

Il metodo di Lobačevskij.

Uno dei metodi più utilizzati per il calcolo delle radici, soprattutto delle radici complesse, è il metodo [14] proposto da N. I. Lobačevskij nel suo libro "Algebra", pubblicato nel 1834. L'idea di base di questo metodo risale a Bernoulli.

Notiamo innanzitutto che se ci viene dato un polinomio le cui radici sono $x_1, x_2, \cdots, x_n$, allora è facile scrivere il polinomio, anch'esso di grado n, le cui radici sono $x^2_1, x^2_2, \cdots, x^2_n$, cioè i quadrati delle radici del polinomio dato. Infatti, se $x_1, x_2, \cdots, x_n$ sono le radici del polinomio

$$x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n ,$$

allora può essere scritto come

$$(x - x_1)(x - x_2) \cdots (x - x_n),$$

e il polinomio

$$x^n - a_1x^{n-1} + a_2x^{n-2} - \cdots \pm a_n ,$$

le cui radici sono le radici del polinomio dato prese con segno opposto, si può scrivere come

$$(x + x_1)(x + x_2) \cdots (x + x_n).$$

Il prodotto di questi due polinomi è quindi:

$$(x^2 - x_1^2)(x^2 - x_2^2) \cdots (x^2 - x_n^2)$$

e quindi contiene solo potenze pari di x. Ponendo $x^2 = y$, otteniamo un polinomio di grado n in y

$$y^n + b_1y^{n-1} + b_2y^{n-2} + \cdots + b_n ,$$

che può essere scritto come

$$(y - x_1^2)(y - x_2^2) \cdots (y - x_n^2),$$

poiché le sue radici sono $x_1^2, x_2^2, \cdots, x_n^2$. Invece di moltiplicare direttamente il polinomio

$$x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n$$

per il polinomio

$$x^n - a_1x^{n-1} + a_2x^{n-2} - \cdots \pm a_n ,$$

possiamo ricavare i coefficienti b_k secondo il seguente schema. Nella prima riga sopra una linea orizzontale scriviamo 1, $a_1, a_2, \cdots, a_n$, e poi sotto la linea, sotto ciascuno di questi coefficienti a_k, scriviamo prima il suo quadrato a_k^2, quindi meno il doppio del prodotto dei suoi vicini

$$- 2a_{k-1}\, a_{k+1} ,$$

quindi più il doppio del prodotto dei coefficienti

$$+ 2a_{k-2}\, a_{k+2} ,$$

simmetrici rispetto a a_k, ecc., alternando di segno fino a quando tutti gli ulteriori coefficienti da una parte o dall'altra sono uguali a zero. I coefficienti b_k, si ottengono quindi sommando ciascuna colonna sotto la riga e moltiplicando il risultato per $(-1)^k$.

Dopo aver ottenuto questi coefficienti 1, b_1, b_2, ⋯, b_n del polinomio le cui radici sono $x_1^2, x_2^2, \cdots, x_n^2$, costruiamo successivamente i coefficienti 1, c_1, c_2, ⋯, c_n del polinomio le cui radici sono i quadrati delle radici del polinomio

$$y^n + b_1 y^{n-1} + b_2 y^{n-2} + \cdots + b_n ,$$

cioè $x_1^4, x_2^4, \cdots, x_n^4$. Allora analogamente otteniamo i coefficienti 1, d_1, d_2, ⋯, d_n del polinomio le cui radici sono $x_1^8, x_2^8, \cdots, x_n^8$; e poi del polinomio le cui radici sono $x_1^{16}, x_2^{16}, \cdots, x_n^{16}$,, e così via.

Consideriamo solo l'idea fondamentale del metodo di Lobačevskij; inoltre ci limitiamo per semplicità al caso in cui tutte le radici dell'equazione siano reali e distinte in valore assoluto. Sia

$$| x_1 | > | x_2 | > \cdots > | x_n |,$$

ovvero, sia x_1 la radice più grande in valore assoluto, x_2 la radice successiva e così via. Sia N un numero sufficientemente grande e sia il polinomio

$$X^n + A_1 X^{n-1} + A_2 X^{n-2} + \cdots + A_n$$

abbia radici pari alla N-esima potenza delle radici $x_1, x_2, \cdots, x_n$ del polinomio dato, cioè,

$$\begin{aligned} -A_1 &= x_1^N + x_2^N + \cdots + x_n^N, \\ A_2 &= x_1^N x_2^N + x_1^N x_3^N + \cdots + x_{n-1}^N x_n^N, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \pm A_n &= x_1^N x_2^N \cdots x_n^N. \end{aligned}$$

Quindi nella successione dei numeri $| x_1^N |, | x_2^N |, \cdots, | x_n^N |$ per N grande ogni numero successivo è tanto più piccolo del suo predecessore che in queste espressioni per $A_1, A_2, \cdots, A_n$, possiamo conservare solo il primo addendo, trascurando la somma di tutti gli addendi rimanenti rispetto al primo. Otteniamo così le formule approssimate

$$x_1^N \approx -A_1, \qquad x_1^N x_2^N \approx A_2,$$

$$x_1^N x_2^N x_3^N \approx A_3, \cdots, x_1^N x_2^N x_3^N \cdots x_n^N \approx \pm A_n,$$

oppure, dividendo a coppie e prendendo le radici N-esime, abbiamo le seguenti formule per x_k:

$$x_1 = \sqrt[N]{-A_1}\,,\ x_2 = \sqrt[N]{-\frac{A_2}{A_1}}\,,\ x_3 = \sqrt[N]{-\frac{A_3}{A_2}}\,,\ \cdots,\ x_n = \sqrt[N]{-\frac{A_n}{A_{n-1}}}\,.$$

Si può dimostrare che è sufficiente estendere il calcolo fino al polinomio i cui coefficienti presi con i segni + - + - ⋯ saranno uguali, con il necessario grado di esattezza, ai quadrati dei corrispondenti coefficienti del polinomio precedente.

Un'esposizione dettagliata del metodo di Lobačevskij si trova nel famoso libro dell'accademico A. N. Krylov "Lezioni sui calcoli approssimati".

Letture consigliate

E. Artin, Galois theory, 2nd ed., University of Notre Dame, 1944.

--, Geometric algebra, Interscience, New York, 1957.

G. Birkhoff and S. MacLane, A survey of modern algebra, 2nd ed., Macmillan, New York, 1953.

F. Klein, Famous problems of elementary geometry. The duplication of the cube, the trisection of an angle, the quadrature of the circle, Dover, New York, 1956.

--, Lectures on the icosahedron and the solution of equations of the fifth degree, Dover, New York, 1956.

M. Marden, The geometry of the zeros of a polynomial in a complex variable, American Mathematical Society, Providence, R. I., 1949.

Note al capitolo IV

[1] Vedi Capitolo XX.

[2] Vedi Capitolo XVII.

[3] Assumiamo che tutti i termini dell'equazione siano trasferiti al membro sinistro e che l'equazione sia divisa per il coefficiente della massima potenza dell'incognita.

[4] La dimostrazione del teorema fondamentale dell'algebra è difficile ed è stata data molto più tardi. Ad esso dedichiamo il §3. Ma la sua validità fu presupposta molto prima che fosse rigorosamente dimostrata.

[5] Cioè un numero complesso che, elevato alla potenza *n*-esima è uguale a uno. Ad esempio, le radici cubiche dell'unità possono avere i valori

$$1, \quad -\frac{1}{2}+\frac{\sqrt{3}}{2}i, \quad -\frac{1}{2}-\frac{\sqrt{3}}{2}i,$$

dove $i = \sqrt{-1}$ (vedi §3). Infatti,

$$\left(-\frac{1}{2}+\frac{\sqrt{3}}{2}i\right)^3 = -\frac{1}{8}-\frac{3}{8}\sqrt{3}\,i+\frac{9}{8}+\frac{3\sqrt{3}}{8}i = 1,$$

e analogamente

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

[6] Dei gruppi di Galois si dirà di più nel §5, Capitolo XX.

[7] Il punto è che esistono equazioni non algebriche, ad esempio $a^z = o$, che sicuramente non hanno radici, né reali né complesse.

[8] Vedi capitolo IX.

[9] Vedi capitolo IX.

[10] Questo teorema è la forma più semplice del teorema del valore medio, menzionato nel capitolo II.

[11] Potremmo dare un'altra prova diretta, senza derivate, ma sarebbe un po' più lunga.

[12] Notiamo che questa affermazione è corretta anche nel caso in cui alcune radici di $f(x)$ siano complesse.

[13] Da queste formule si ottiene anche una dimostrazione rigorosa delle due asserzioni fatte dall'esame dei diagrammi. Cioè i valori α_n, (o β_n) al crescere di n cambiano monotonicamente, ad esempio nel caso I essi aumentano e sono limitati, cioè, in virtù del lemma di Weierstrass, si avvicinano a qualche limite α. Sostituendo α_n in queste formule con il suo limite α, otteniamo $\alpha = \alpha - [f(\alpha)/f'(\alpha)]$ da cui $f(\alpha) = 0$, ovvero α è una radice di f.

[14] Questo metodo fu scoperto indipendentemente da Dandelin (1826), N. I. Lobačevskij (1834) e Graeffe (1837).

Nota editoriale

Sono stati corretti errori di trascrizione di indici, esponenti, lettere greche, segni e derivate. La figura 1 e la formula della rappresentazione trigonometrica, scambiate nella copia elettronica, sono ricollocate correttamente. Sono inoltre esplicitati o corretti alcuni punti del testo di partenza: gli interi nell'esempio di Fermat sono non nulli; nel teorema sul minimo si considera $|z| > R$ e, nel passaggio al limite, la differenza $|f(z_n)| - |f(z_0)|$; nella dimostrazione della legge di Cartesio sono resi coerenti gli indici dell'ultimo coefficiente e la potenza di x che raccoglie le radici nulle; nell'osservazione successiva si legge «radici reali»; nella discussione del criterio di Hurwitz si contano n mezzi giri. Nel metodo di Lobačevskij è eliminato un 1 estraneo all'elenco delle radici e si esplicita il fattore $(-1)^k$ per ottenere b_k con i segni della forma monica adottata. La formula nella dimostrazione della legge di Cartesio è ricomposta tipograficamente.

La successione di Sturm termina in una costante quando il polinomio è privo di radici multiple; in caso contrario si applica la costruzione dopo aver diviso il polinomio per il massimo comune divisore con la sua derivata. Nel conteggio delle variazioni di segno si omettono gli zeri. La formulazione storica e le valutazioni sullo stato dei problemi sono conservate come nell'originale.

CAPITOLO V

Equazioni differenziali ordinarie

I. G. Petrovskij

§ 1. Introduzione

Esempi di equazioni differenziali.

Le equazioni incontrate finora riguardavano per lo più la ricerca del valore numerico di una grandezza. Quando, ad esempio, cercando i massimi e i minimi di una funzione, risolvevamo un'equazione per trovare i punti nei quali si annulla la sua velocità di variazione, oppure quando, nel capitolo IV, cercavamo le radici di un polinomio, in entrambi i casi si trattava di determinare singoli numeri. Nelle applicazioni della matematica sorgono però spesso problemi di natura qualitativamente diversa, nei quali l'incognita è essa stessa una funzione, una legge che esprime la dipendenza di alcune variabili da altre. Studiando il raffreddamento di un corpo, ad esempio, dobbiamo stabilire come varia nel tempo la sua temperatura; per descrivere il moto di un pianeta o di una stella dobbiamo determinare come le sue coordinate dipendono dal tempo, e così via.

Molto spesso possiamo costruire un'equazione per determinare le funzioni incognite cercate: tali equazioni si chiamano equazioni funzionali. La loro natura può essere estremamente varia; anzi, possiamo dire di averne già incontrato le forme più semplici ed elementari quando abbiamo considerato le funzioni implicite.

Il problema della determinazione di funzioni incognite ci occuperà nei capitoli V, VI e VII. In questo capitolo e nel successivo considereremo la classe più importante di equazioni che servono a determinarle: le equazioni differenziali, nelle quali compaiono, oltre alla funzione incognita, anche le sue derivate di vari ordini. Ecco alcuni esempi:

$$\frac{dx}{dt} + P(t)\,x = Q(t), \frac{d^2x}{dt^2} + m^2x = A\sin\omega t, \frac{d^2x}{dt^2} = tx,$$

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \qquad (1)$$

Nelle prime tre equazioni la funzione incognita è indicata con la lettera x e la variabile indipendente con t; nelle ultime tre l'incognita è indicata con u e dipende da due argomenti, x e t, oppure x e y.

La grande importanza delle equazioni differenziali nella matematica, e soprattutto nelle sue applicazioni, deriva principalmente dal fatto che lo studio di molti problemi della fisica e della tecnica si può ricondurre alla loro risoluzione.

I calcoli necessari alla costruzione di macchine elettriche o di dispositivi radiotecnici, il calcolo delle traiettorie dei proiettili, lo studio della stabilità di un aeroplano in volo o dell'andamento di una reazione chimica dipendono tutti dalla risoluzione di equazioni differenziali.

Spesso le leggi fisiche che governano un fenomeno sono scritte sotto forma di equazioni differenziali, che ne forniscono così un'espressione quantitativa, numerica, esatta. Nei capitoli successivi il lettore vedrà come si esprimono in questa forma le leggi di conservazione della massa e dell'energia termica. Le leggi della meccanica scoperte da Newton permettono di studiare mediante equazioni differenziali il comportamento di qualunque sistema meccanico.

Illustriamo questo fatto con un semplice esempio. Consideriamo un punto materiale di massa m che si muove lungo un asse Ox, e indichiamo con x la sua coordinata all'istante t. La coordinata varia nel tempo: conoscere l'intero moto del punto equivale a conoscere la dipendenza funzionale di x da t. Supponiamo che il moto sia prodotto da una forza F, il cui valore dipenda dalla posizione del punto, definita dalla coordinata x, dalla velocità $v = dx/dt$ e dal tempo t: dunque $F = F(x, dx/dt, t)$. Secondo le leggi della meccanica, la forza produce un'accelerazione $w = d^2x/dt^2$ tale che il prodotto di w per la massa m è uguale alla forza. In ogni istante del moto vale quindi l'equazione

$$m\frac{d^2x}{dt^2} = F\left(x, \frac{dx}{dt}, t\right). \qquad (2)$$

Questa è l'equazione differenziale che deve soddisfare la funzione $x(t)$ che descrive il moto del punto. Essa rappresenta semplicemente le leggi della meccanica. La sua importanza consiste nel permetterci di ricondurre il problema meccanico della determinazione del moto al problema matematico della risoluzione di un'equazione differenziale.

Più avanti il lettore troverà altri esempi di come lo studio di diversi processi fisici si possa ricondurre all'indagine delle equazioni differenziali.

La teoria delle equazioni differenziali cominciò a svilupparsi alla fine del Seicento, quasi contemporaneamente alla nascita del calcolo differenziale e integrale. Oggi queste equazioni sono diventate uno strumento potente per lo studio dei fenomeni naturali e hanno consentito progressi immensi nella meccanica, nell'astronomia, nella fisica e nella tecnica. Studiando le equazioni differenziali del moto dei corpi celesti, Newton dedusse le leggi del moto planetario che Keplero aveva scoperto empiricamente. Nel 1846 Le Verrier previde l'esistenza del pianeta Nettuno e ne determinò la posizione nel cielo sulla base di un'analisi numerica delle stesse equazioni.

Per descrivere in termini generali i problemi della teoria, osserviamo anzitutto che un'equazione differenziale ha, in generale, non una sola soluzione, ma infinite: esiste cioè un insieme infinito di funzioni che la soddisfano. L'equazione del moto di un punto materiale, ad esempio, deve essere soddisfatta da ogni moto prodotto dalla forza assegnata $F(x, dx/dt, t)$, indipendentemente dalla posizione di partenza e dalla velocità iniziale. A ogni singolo moto corrisponde una particolare dipendenza di x dal tempo t. Poiché sotto l'azione di una stessa forza possono aversi infiniti moti, l'equazione differenziale (2) possiede un insieme infinito di soluzioni.

Ogni equazione differenziale definisce dunque, in generale, un'intera classe di funzioni che la soddisfano. Il problema fondamentale della teoria consiste nello studio di queste funzioni. La teoria deve permetterci di formarci un'idea sufficientemente ampia delle proprietà di tutte le funzioni che soddisfano l'equazione: esigenza particolarmente importante nelle applicazioni alle scienze naturali. Deve inoltre fornire i mezzi per determinarne i valori numerici quando occorrano per un calcolo. Diremo più avanti come si possono trovare tali valori.

Se la funzione incognita dipende da un solo argomento, l'equazione si chiama equazione differenziale ordinaria. Se invece dipende da più argomenti e nell'equazione compaiono derivate rispetto ad alcuni o a tutti questi argomenti, si parla di equazione differenziale alle derivate parziali. Le prime tre equazioni della (1) sono ordinarie; le ultime tre sono alle derivate parziali.

La teoria delle equazioni alle derivate parziali presenta numerose peculiarità che la distinguono sostanzialmente da quella delle equazioni ordinarie. Le sue idee fondamentali saranno esposte nel prossimo capitolo; qui esamineremo soltanto le equazioni differenziali ordinarie.

Consideriamo alcuni esempi.

Esempio 1.

La legge del decadimento del radio afferma che la velocità di decadimento è proporzionale alla quantità di radio presente. Supponiamo di sapere che all'istante $t = t_0$ vi erano R_0 grammi di radio. Vogliamo determinare la quantità presente a un qualunque istante successivo t.

Indichiamo con $R(t)$ la quantità di radio non ancora decaduto all'istante t. La velocità di decadimento è $-dR/dt$. Poiché essa è proporzionale a R, abbiamo

$$-\frac{dR}{dt} = kR, \qquad (3)$$

dove k è una costante.

Per risolvere il problema dobbiamo determinare una funzione a partire dall'equazione differenziale (3). A questo scopo osserviamo che la funzione inversa di $R(t)$ soddisfa l'equazione

$$-\frac{dt}{dR} = \frac{1}{kR}, \qquad (4)$$

poiché $dt/dR = 1/(dR/dt)$. Dal calcolo integrale sappiamo che l'equazione (4) è soddisfatta da ogni funzione della forma

$$t = -\frac{1}{k}\ln R + C,$$

dove C è una costante arbitraria. Da questa relazione ricaviamo R in funzione di t. Otteniamo

$$R = e^{-kt+kC} = C_1 e^{-kt}. \qquad (5)$$

Tra tutte le soluzioni (5) dell'equazione (3) dobbiamo scegliere quella che per $t = t_0$ assume il valore R_0. La si ottiene ponendo $C_1 = R_0 e^{kt_0}$.

Dal punto di vista matematico, l'equazione (3) esprime una legge semplicissima di variazione nel tempo della funzione R: la velocità di diminuzione $-dR/dt$ è proporzionale al valore della funzione stessa. Questa legge è soddisfatta non soltanto dal decadimento radioattivo, ma anche da molti altri fenomeni fisici.

La stessa legge si incontra, ad esempio, nello studio del raffreddamento di un corpo: la velocità con cui diminuisce la sua quantità di calore è proporzionale alla differenza tra la temperatura del corpo e quella dell'ambiente circostante. Essa compare anche in molti altri processi fisici. Il campo di applicazione dell'equazione (3) è quindi molto più vasto del particolare problema del decadimento radioattivo dal quale l'abbiamo ricavata.

Esempio 2.

Un punto materiale di massa m si muova lungo l'asse orizzontale Ox in un mezzo resistente, per esempio un liquido o un gas, sotto l'azione della forza elastica di due molle (figura 1). Secondo la legge di Hooke, tale forza è diretta verso la posizione di equilibrio ed è proporzionale allo scostamento da essa. Poniamo la posizione di equilibrio nel punto $x = 0$. La forza elastica è allora $-bx$, dove $b > 0$.

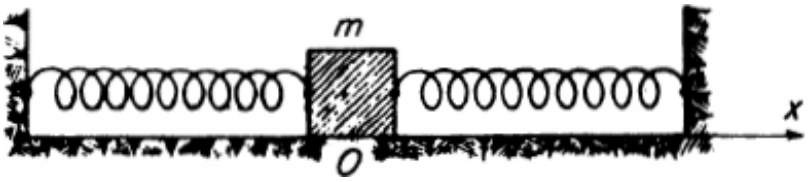


Figura 1.

Supponiamo che la resistenza del mezzo sia proporzionale alla velocità del moto, cioè uguale a $-a(dx/dt)$, con $a > 0$. Il segno meno indica che il mezzo si oppone al moto. Questa ipotesi sulla resistenza è confermata dall'esperienza.

Per la legge fondamentale di Newton, il prodotto della massa del punto materiale per la sua accelerazione è uguale alla somma delle forze agenti su di esso. Pertanto

$$m\frac{d^2x}{dt^2} = -bx - a\frac{dx}{dt}. \qquad (6)$$

La funzione $x(t)$, che descrive la posizione del punto mobile in ogni istante t, soddisfa dunque l'equazione differenziale (6). Ne studieremo le soluzioni in una delle sezioni successive.

Se, oltre alle forze considerate, sul punto agisce un'ulteriore forza F, esterna al sistema, l'equazione del moto (6) assume la forma

$$m\frac{d^2x}{dt^2} = -bx - a\frac{dx}{dt} + F \qquad (6')$$

Esempio 3.

Un pendolo matematico è un punto materiale di massa m sospeso a un filo di lunghezza l. Supponiamo che il pendolo rimanga sempre in uno stesso piano, quello della figura 2. La forza che tende a ricondurlo nella posizione verticale OA è la forza di gravità mg agente sul punto materiale. La posizione all'istante t è determinata dall'angolo ϕ di scostamento dalla verticale OA; scegliamo come verso positivo quello antiorario. L'arco $AA' = l\phi$ misura lo spostamento del punto dalla posizione di equilibrio A. La velocità v è diretta lungo la tangente alla circonferenza e ha il valore algebrico

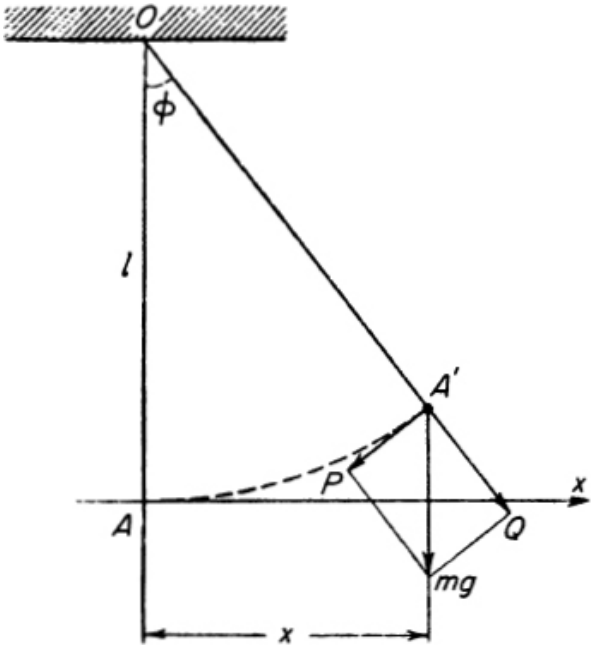


Figura 2.

$$v = l\frac{d\phi}{dt}.$$

Per stabilire l'equazione del moto scomponiamo la forza di gravità mg nelle componenti Q e P, dirette rispettivamente lungo il raggio OA′ e lungo la tangente alla circonferenza. La componente Q non modifica il valore della velocità v: la reazione del vincolo OA′ impedisce lo spostamento radiale. Solo la componente P può modificarlo. Essa è sempre diretta verso la posizione di equilibrio A, e quindi tende a diminuire ϕ se l'angolo è positivo e ad aumentarlo se è negativo. Il suo valore algebrico è $-mg \sin \phi$; l'equazione del moto del pendolo è dunque

$$m\frac{dv}{dt} = -mg \sin \phi$$

ossia

$$\frac{d^2\phi}{dt^2} = -\frac{g}{l}\sin\phi. \qquad (7)$$

È interessante osservare che le soluzioni di questa equazione non si possono esprimere mediante una combinazione finita di funzioni elementari. L'insieme di queste funzioni è troppo ristretto per descrivere esattamente persino un processo fisico semplice come l'oscillazione di un pendolo matematico. Vedremo che le equazioni differenziali risolvibili mediante funzioni elementari non sono molto numerose. Accade perciò assai spesso che lo studio di un'equazione incontrata nella fisica o nella meccanica ci induca a introdurre e studiare nuove classi di funzioni, ampliando così l'insieme degli strumenti disponibili per risolvere problemi applicativi.

Limitiamoci ora alle piccole oscillazioni del pendolo, per le quali possiamo supporre, con un piccolo errore, che l'arco AA′ coincida con la sua proiezione x sull'asse orizzontale Ox e che sin ϕ sia uguale a ϕ. Allora $\phi \approx \sin\phi = x/l$, e l'equazione del moto assume la forma più semplice

$$\frac{d^2x}{dt^2} = -\frac{g}{l}x. \qquad (8)$$

Vedremo che questa equazione si risolve mediante funzioni trigonometriche, con le quali possiamo descrivere con sufficiente precisione le «piccole oscillazioni» del pendolo.

Esempio 4.

Il risonatore acustico di Helmholtz (figura 3) consiste in un recipiente V pieno d'aria, di volume v, munito di un collo cilindrico F. In prima approssimazione possiamo considerare l'aria contenuta nel collo come un tappo di massa $m = \rho sl$, dove ρ è la densità dell'aria, s l'area della sezione trasversale del collo e l la sua lunghezza. Se questa massa d'aria si sposta di una quantità x dalla posizione di equilibrio, la pressione dell'aria nel recipiente di volume v cambia rispetto al valore iniziale p di una quantità che indicheremo con Δp.

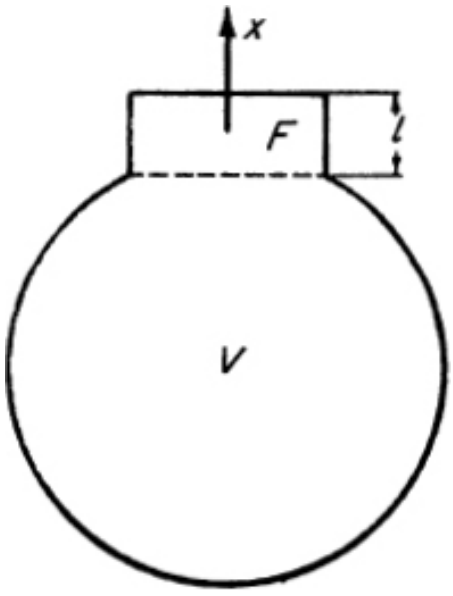


Figura 3.

$$m = \rho s l, \qquad (9)$$

Supponiamo che la pressione p e il volume v soddisfino la legge adiabatica $pv^k = C$. Trascurando le grandezze di ordine superiore, abbiamo

$$\Delta p \cdot v^k + pkv^{k-1} \cdot \Delta v = 0$$

e

$$\Delta p = -kp \frac{\Delta v}{v} = -\frac{kps}{v} x. \qquad (10)$$

Nel nostro caso $\Delta v = sx$. L'equazione del moto della massa d'aria nel collo si può scrivere così:

$$m \frac{d^2x}{dt^2} = \Delta p \cdot s. \qquad (11)$$

Qui $\Delta p \cdot s$ è la forza esercitata dal gas nel recipiente sulla colonna d'aria del collo. Dalle (10) e (11) ricaviamo

$$\rho l \frac{d^2x}{dt^2} = -\frac{kps}{v} x, \qquad (12)$$

dove ρ, p, v, l, k e s sono costanti.

Esempio 5.

Un'equazione della forma (6) compare anche nello studio delle oscillazioni elettriche in un semplice circuito oscillante, schematizzato nella figura 4. A sinistra troviamo un condensatore di capacità C, collegato in serie con una bobina di induttanza L e una resistenza R. Supponiamo che a un certo istante vi sia una tensione ai capi del condensatore. In assenza di induttanza, la corrente fluirebbe finché i suoi terminali non raggiungessero lo stesso potenziale. La presenza dell'induttanza modifica la situazione, perché nel circuito si generano oscillazioni elettriche. Per determinarne la legge indichiamo con $v(t)$, o semplicemente con v, la tensione ai capi del condensatore all'istante t, con $I(t)$ la corrente e con R la resistenza. Per note leggi fisiche, $I(t)R$ è sempre uguale alla forza elettromotrice totale, somma, con i segni indicati nella (13), del contributo del condensatore e della forza elettromotrice di autoinduzione $-L(dI/dt)$. Quindi

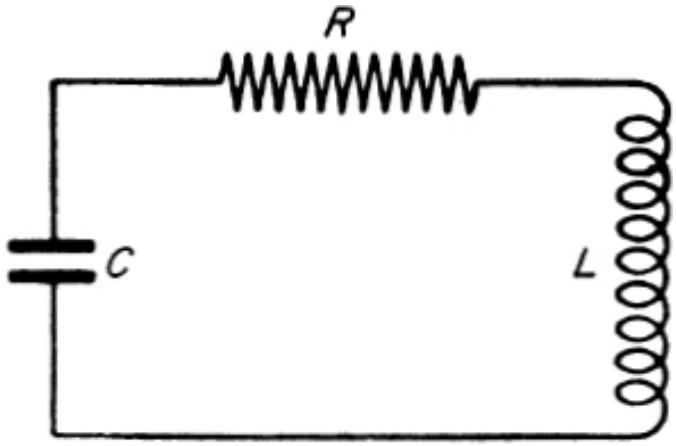


Figura 4.

$$IR = -v - L\frac{dI}{dt}. \qquad (13)$$

Indichiamo con $Q(t)$ la carica del condensatore all'istante t. La corrente nel circuito è allora $I = dQ/dt$. Poiché la differenza di potenziale è $v(t) = Q(t)/C$, si ha $I = C(dv/dt)$ e la (13) si trasforma in

$$LC\frac{d^2v}{dt^2} + RC\frac{dv}{dt} + v = 0. \qquad (14)$$

Esempio 6.

La figura 5 mostra lo schema di un generatore a valvola elettronica di oscillazioni elettromagnetiche. Il sistema oscillante fondamentale è il circuito formato dalla capacità C, dalla resistenza R e dall'induttanza L. La bobina L' e la valvola al centro della figura realizzano la cosiddetta «retroazione»: collegano la sorgente di energia, cioè la batteria B, al circuito L-R-C. K è il catodo della valvola, A l'anodo e S la griglia. Nel circuito possono così instaurarsi «autooscillazioni». In ogni sistema reale che oscilla, l'energia viene trasformata in calore o dissipata in altra forma verso i corpi circostanti; per mantenere un regime oscillatorio stazionario occorre quindi una sorgente esterna di energia. Le autooscillazioni si distinguono dagli altri processi oscillatori perché, per mantenere tale regime, la sorgente esterna non deve essere periodica. Il sistema è costruito in modo che una sorgente costante, nel nostro caso la batteria B, mantenga le oscillazioni stazionarie. Sono esempi di sistemi autooscillanti un orologio, un campanello elettrico, una corda con l'archetto mosso dalla mano del musicista, l'apparato vocale umano, e così via.

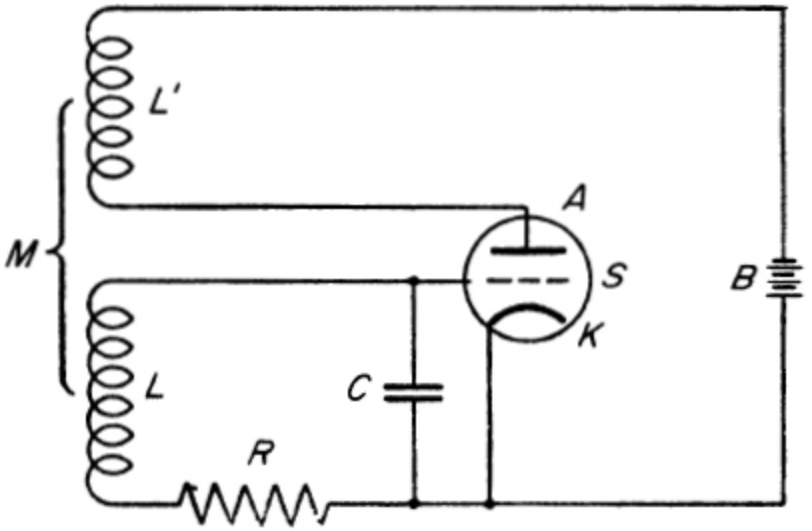


Figura 5.

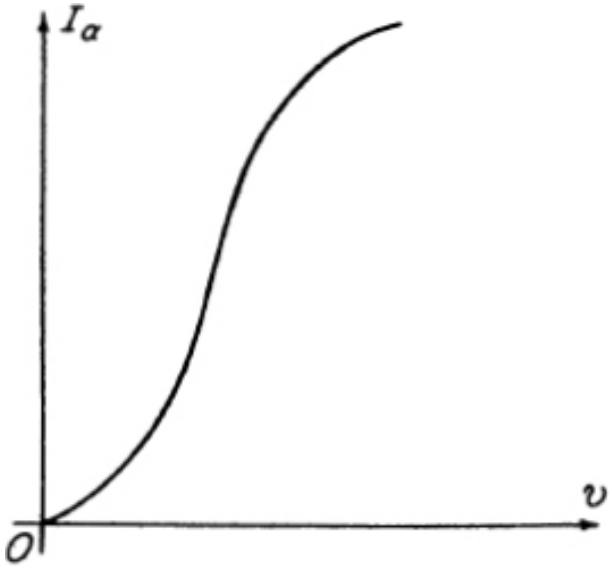


Figura 6.

La corrente $I(t)$ nel circuito oscillante L-R-C soddisfa l'equazione

$$L\frac{dI}{dt} + RI + v = M\frac{dI_a}{dt}. \qquad (15)$$

Qui $v = v(t)$ è la tensione ai capi del condensatore all'istante t, $I_a(t)$ è la corrente anodica che attraversa la bobina L' e M è il coefficiente di mutua induzione tra le bobine L e L'. Rispetto alla (13), l'equazione (15) contiene il termine aggiuntivo $M(dI_a/dt)$.

Supponiamo che la corrente anodica $I_a(t)$ dipenda soltanto dalla tensione tra la griglia S e il catodo della valvola, trascurando cioè la reazione dell'anodo, e che questa tensione sia uguale a $v(t)$, ai capi del condensatore C. L'andamento della dipendenza funzionale di I_a da v è illustrato nella figura 6. La curva viene solitamente approssimata con una parabola cubica, la cui equazione è

$$I_a = a_1 v + a_2 v^2 + a_3 v^3.$$

Sostituendo questa espressione nel secondo membro della (15) e tenendo conto che

$$\frac{dv}{dt} = I,$$

otteniamo per v l'equazione

$$L\frac{d^2v}{dt^2} + [R - M(a_1 + 2a_2 v + 3a_3 v^2)]\frac{dv}{dt} + v = 0. \qquad (16)$$

Negli esempi considerati, la ricerca di determinate grandezze caratteristiche di un processo fisico si riconduce alla ricerca delle soluzioni di equazioni differenziali ordinarie.

Problemi della teoria delle equazioni differenziali.

Diamo ora definizioni precise. Un'equazione differenziale ordinaria di ordine n in una funzione incognita y è una relazione della forma

$$F[x, y(x), y'(x), y''(x), \cdots, y^{(n)}(x)] = 0 \qquad (17)$$

tra la variabile indipendente x e le grandezze

$$y(x), y'(x) = \frac{dy}{dx}, y''(x) = \frac{d^2y}{dx^2}, \cdots, y^{(n)}(x) = \frac{d^n y}{dx^n}.$$

L'ordine di un'equazione differenziale è l'ordine della derivata più alta della funzione incognita che compare nell'equazione. L'equazione dell'esempio 1 è quindi del primo ordine, mentre quelle degli esempi 2, 3, 4, 5 e 6 sono del secondo ordine.

Una funzione $\phi(x)$ si chiama soluzione dell'equazione differenziale (17) se, sostituendo $\phi(x)$ a y, $\phi'(x)$ a y', $\cdots$, $\phi^{(n)}(x)$ a $y^{(n)}$, si ottiene un'identità.

I problemi della fisica e della tecnica conducono spesso a sistemi di equazioni differenziali ordinarie con più funzioni incognite, tutte dipendenti dallo stesso argomento, e con le loro derivate rispetto a quell'argomento.

Per rendere più concreta l'esposizione, ci occuperemo soprattutto di una sola equazione differenziale ordinaria, di ordine non superiore al secondo e con una sola funzione incognita. Attraverso questo esempio si possono chiarire le proprietà essenziali di tutte le equazioni differenziali ordinarie e dei sistemi nei quali il numero delle funzioni incognite è uguale al numero delle equazioni.

Abbiamo già osservato che, di regola, ogni equazione differenziale possiede un insieme infinito di soluzioni. Illustriamolo anzitutto con considerazioni intuitive basate sugli esempi 2-6. In ciascuno di essi l'equazione è già completamente determinata dalla struttura fisica del sistema; eppure, in ogni sistema sono possibili molti moti differenti. È evidente, ad esempio, che il pendolo descritto dalla (8) può oscillare con ampiezze diverse. A ogni oscillazione

corrisponde una diversa soluzione della (8), che deve quindi averne infinite. Si può dimostrare che la (8) è soddisfatta da ogni funzione della forma

$$x = C_1 \cos \sqrt{\frac{g}{l}}\, t + C_2 \sin \sqrt{\frac{g}{l}}\, t, \qquad (18)$$

dove C_1 e C_2 sono costanti arbitrarie.

È altrettanto chiaro dal punto di vista fisico che il moto del pendolo è completamente determinato soltanto se, a un certo istante t_0, sono assegnati il valore iniziale x_0 di x, cioè lo spostamento iniziale del punto dalla posizione di equilibrio, e la velocità iniziale $x_0' = (dx/dt)\,|_{t=0}$. Queste condizioni iniziali determinano le costanti C_1 e C_2 nella formula (18).

Allo stesso modo, anche le equazioni differenziali ricavate negli altri esempi possiedono infinite soluzioni.

In generale, sotto ipotesi molto ampie sull'equazione differenziale (17), di ordine n e con una sola funzione incognita, si può dimostrare che essa ha infinite soluzioni. Più precisamente: se, per un certo «valore iniziale» dell'argomento, assegniamo un «valore iniziale» alla funzione incognita e a tutte le sue derivate fino all'ordine $n - 1$, si può trovare una soluzione della (17) che assuma questi valori. Si può inoltre dimostrare che tali condizioni iniziali determinano completamente la soluzione: ne esiste una sola che le soddisfa. Torneremo più dettagliatamente sulla questione. Per ora è essenziale osservare che i valori iniziali della funzione e delle sue prime $n - 1$ derivate possono essere assegnati arbitrariamente: possiamo scegliere liberamente gli n valori che definiscono uno «stato iniziale» per la soluzione cercata.

Se vogliamo costruire una formula che comprenda, possibilmente, tutte le soluzioni di un'equazione differenziale di ordine n, essa deve contenere n costanti arbitrarie indipendenti, che permettano di imporre n condizioni iniziali. Le soluzioni espresse in questa forma si chiamano solitamente soluzioni generali dell'equazione. Ad esempio, una soluzione generale della (8) è la formula (18), che contiene due costanti arbitrarie; una soluzione generale della (3) è data dalla formula (5).

Cerchiamo ora di delineare, nelle loro linee più generali, i problemi della teoria delle equazioni differenziali. Sono numerosi e diversi; indicheremo soltanto i più importanti.

Quando l'equazione è assegnata insieme alle condizioni iniziali, la sua soluzione è completamente determinata. Costruire formule che la esprimano esplicitamente è uno dei primi problemi della teoria. Ciò è possibile soltanto in casi semplici; ma quando si trovano, queste formule sono di grande aiuto per il calcolo e per lo studio della soluzione.

La teoria deve permetterci di comprendere il comportamento di una soluzione: se sia monotona oppure oscillante, se sia periodica o si avvicini a una funzione periodica, e così via.

Supponiamo di modificare i valori iniziali della funzione incognita e delle sue derivate, cioè lo stato iniziale del sistema fisico. Cambierà anche la soluzione, perché l'intero processo fisico si svolgerà diversamente. La teoria deve permetterci di valutare questo cambiamento. In particolare, a piccole variazioni dei valori iniziali corrisponderanno piccole variazioni della

soluzione, che sarà quindi stabile sotto questo aspetto? Oppure piccoli cambiamenti delle condizioni iniziali potranno provocare grandi variazioni della soluzione, rendendola instabile?

Dobbiamo inoltre poter costruire un quadro qualitativo e, quando possibile, quantitativo del comportamento non soltanto delle singole soluzioni, ma anche del loro insieme.

Nella costruzione delle macchine si presenta spesso il problema di scegliere i parametri di un apparecchio in modo da garantirne il buon funzionamento. Questi parametri compaiono come determinate grandezze nell'equazione differenziale corrispondente. La teoria deve aiutarci a chiarire che cosa accade alle soluzioni, e quindi al funzionamento dell'apparecchio, quando modifichiamo l'equazione, cioè i parametri dell'apparecchio.

Infine, quando occorre eseguire un calcolo, dobbiamo determinare numericamente la soluzione. La teoria deve allora fornire all'ingegnere e al fisico i metodi più rapidi ed economici per calcolarla.

§ 2. Equazioni differenziali lineari a coefficienti costanti

Per alcune importanti classi di equazioni differenziali ordinarie, la soluzione generale si può esprimere mediante funzioni semplici e ben note. Una di queste classi è costituita dalle equazioni differenziali a coefficienti costanti che sono lineari rispetto alla funzione incognita e alle sue derivate, dette brevemente lineari. Ne sono esempi le equazioni (3), (6), (8) e (14). Un'equazione lineare si chiama omogenea se non contiene un termine indipendente dalla funzione incognita; in caso contrario si chiama non omogenea.

Equazioni lineari omogenee del secondo ordine a coefficienti costanti.

Queste equazioni hanno la forma

$$m\frac{d^2x}{dt^2} + a\frac{dx}{dt} + bx = 0, \qquad (6)$$

dove m, a e b sono costanti. Supporremo m positivo: ciò non limita la generalità, perché, se necessario, possiamo cambiare il segno di tutti i coefficienti, purché $m \neq 0$, come assumiamo.

Cerchiamo una soluzione sotto forma di funzione esponenziale $e^{\lambda t}$ e chiediamoci come scegliere la costante λ affinché $x = e^{\lambda t}$ soddisfi l'equazione. Sostituendo $x = e^{\lambda t}$, $dx/dt = \lambda e^{\lambda t}$ e $d^2x/dt^2 = \lambda^2 e^{\lambda t}$ nel primo membro della (6), otteniamo

$$e^{\lambda t}(m\lambda^2 + a\lambda + b).$$

Affinché $x(t) = e^{\lambda t}$ sia una soluzione della (6), è quindi necessario e sufficiente che

$$m\lambda^2 + a\lambda + b = 0. \qquad (19)$$

Se λ_1 e λ_2 sono due radici reali della (19), è facile dimostrare che ogni funzione della forma

$$x = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}, \qquad (20)$$

è una soluzione della (6), con C_1 e C_2 costanti arbitrarie.

Dimostreremo più avanti che, quando le radici reali della (19) sono distinte, la formula (20) fornisce tutte le soluzioni della (6).

Osserviamo le seguenti importanti proprietà delle soluzioni della (6):

1. La somma di due soluzioni è ancora una soluzione.

2. Il prodotto di una soluzione per una costante è ancora una soluzione.

Se λ_1 è una radice multipla della (19), cioè $m\lambda_1^2 + a\lambda_1 + b = 0$ e $2m\lambda_1 + a = 0$,[1] anche la funzione $te^{\lambda_1 t}$ è una soluzione della (6). Infatti, sostituendo questa funzione e le sue derivate nel primo membro della (6), otteniamo

$$te^{\lambda_1 t}(m\lambda_1^2 + a\lambda_1 + b) + e^{\lambda_1 t}(2m\lambda_1 + a),$$

espressione che, in virtù delle relazioni precedenti, è identicamente nulla.

In questo caso la soluzione generale della (6) ha la forma

$$x = C_1 e^{\lambda_1 t} + C_2 t e^{\lambda_1 t}. \qquad (21)$$

Supponiamo ora che la (19) abbia radici complesse. Esse sono coniugate, perché m, a e b sono reali. Sia $\lambda = \alpha \pm i\beta$. L'equazione

$$m(\alpha + i\beta)^2 + a(\alpha + i\beta) + b = 0$$

equivale alle due equazioni

$$m\alpha^2 - m\beta^2 + a\alpha + b = 0 \text{ and } 2m\alpha\beta + a\beta = 0. \qquad (22)$$

È facile verificare che, in questo caso, $x = e^{\alpha t} \cos \beta t$ e $x = e^{\alpha t} \sin \beta t$ sono soluzioni della (6). Sostituendo, per esempio, la funzione $x(t) = e^{\alpha t} \cos \beta t$ e le sue derivate nel primo membro della (6), troviamo

$$e^{\alpha t} \cos \beta t(m\alpha^2 - m\beta^2 + a\alpha + b) - e^{\alpha t} \sin \beta t(2m\alpha\beta + a\beta).$$

Per le relazioni (22), questa espressione è identicamente nulla.

Quando la (19) ha radici complesse, la soluzione generale della (6) è

$$x = C_1 e^{\alpha t} \sin \beta t + C_2 e^{\alpha t} \cos \beta t, \qquad (23)$$

dove C_1 e C_2 sono costanti arbitrarie.

Conoscendo le radici della (19), chiamata equazione caratteristica, possiamo dunque scrivere la soluzione generale della (6).

Analogamente, la soluzione generale di un'equazione lineare omogenea di ordine n a coefficienti costanti,

$$a_n \frac{d^n x}{dt^n} + a_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \cdots + a_1 \frac{dx}{dt} + a_0 x = 0$$

si può esprimere mediante combinazioni polinomiali di t e di funzioni esponenziali e trigonometriche, purché si conoscano le radici dell'equazione algebrica

$$a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0 = 0,$$

anch'essa chiamata equazione caratteristica. Il problema di integrare un'equazione differenziale ordinaria lineare a coefficienti costanti si riconduce così a un problema algebrico.

Dimostriamo ora che le formule (20), (21) e (23) forniscono tutte le soluzioni della (6). Osserviamo che in esse si possono sempre scegliere C_1 e C_2 in modo che $x(t)$ soddisfi condizioni iniziali arbitrarie $x(t_0) = x_0$, $x'(t_0) = x_0'$. Basta determinare le due costanti dal sistema

$$\begin{aligned} x_0 &= C_1 e^{\lambda_1 t_0} + C_2 e^{\lambda_2 t_0}, \\ x_0' &= \lambda_1 C_1 e^{\lambda_1 t_0} + \lambda_2 C_2 e^{\lambda_2 t_0}. \end{aligned}$$

nel caso della formula (20), oppure da due equazioni analoghe nel caso delle (21) e (23). Se esistesse una soluzione della (6) non compresa tra quelle costruite, esisterebbero due soluzioni distinte con le stesse condizioni iniziali. La loro differenza $x_1(t)$ non sarebbe identicamente nulla e soddisferebbe le condizioni iniziali nulle $x_1(t_0) = 0$, $x_1(t_0) = 0$, $x_1'(t_0) = 0$. Mostreremo invece che una soluzione della (6) con condizioni iniziali nulle può essere soltanto $x_1(t) = 0$. Supponiamo dapprima $m > 0$, $a > 0$ e $b > 0$. Moltiplichiamo entrambi i membri dell'equazione

$$m \frac{d^2 x_1}{dt^2} + a \frac{dx_1}{dt} + b x_1 = 0 \qquad (24)$$

per $2(dx_1/dt)$. Poiché

$$2 \frac{dx_1}{dt} \cdot \frac{d^2 x_1}{dt^2} = \frac{d}{dt} \left(\frac{dx_1}{dt} \right)^2 \quad \text{and} \quad 2 x_1(t) \frac{dx_1}{dt} = \frac{d}{dt} (x_1^2),$$

la (24) si può scrivere nella forma

$$\frac{d}{dt} \left[m \left(\frac{dx_1}{dt} \right)^2 \right] + 2a \left(\frac{dx_1}{dt} \right)^2 + b \frac{d}{dt} (x_1^2) = 0.$$

Integrando questa identità tra t_0 e t, otteniamo

$$m \left(\frac{dx_1}{dt} \right)^2 + 2a \int_{t_0}^{t} \left(\frac{dx_1}{dt} \right)^2 dt + b x_1^2(t) = 0.$$

Per $t \geq t_0$ tutti i termini del primo membro sono non negativi, mentre il secondo membro è nullo: necessariamente $x_1(t) = 0$. La conclusione vale anche prima di t_0, applicando il ragionamento all'equazione con il tempo invertito e usando la trasformazione descritta nel capoverso seguente.

Per estendere il risultato a coefficienti costanti arbitrari m, a e b, consideriamo $y_1(t) = x_1(t)e^{-\alpha t}$, che soddisfa anch'essa condizioni iniziali nulle. Scegliendo $\alpha > 0$ sufficientemente grande, la funzione y_1 soddisfa un'equazione della forma (6) con tutti e tre i coefficienti positivi. La si ricava facilmente sostituendo $x_1(t) = y_1(t)e^{\alpha t}$ e le sue derivate nella (6). Per quanto dimostrato, $y_1(t) \equiv 0$ e quindi anche $x_1(t) = y_1(t)e^{\alpha t} \equiv 0$.

Abbiamo così dimostrato che le formule (20), (21) e (23) danno tutte le soluzioni della (6).

Vediamo quali informazioni queste formule forniscono sul comportamento delle soluzioni. A questo scopo ricordiamo le formule

$$\lambda_{1,2} = -\frac{a}{2m} \pm \sqrt{\frac{a^2}{4m^2} - \frac{b}{m}} \qquad (25)$$

per le radici della (19). In accordo con le applicazioni fisiche che ci hanno condotti alla (6), supporremo $m > 0$, $a \geqslant 0$ e $b > 0$.

Caso 1.

$a^2 > 4bm$. Le due radici dell'equazione caratteristica (19) sono reali, negative e distinte. La soluzione generale della (6) è allora la funzione $x(t)$ della formula (20). Tutte queste funzioni, insieme alle loro derivate prime, tendono a zero per $t \to +\infty$ e, salvo la soluzione identicamente nulla, possono annullarsi al più una volta. Ne segue che $x(t)$ ha al più un massimo o un minimo. Fisicamente, la resistenza del mezzo è tanto grande da impedire le oscillazioni. Il punto mobile non può attraversare la posizione di equilibrio $x = 0$ più di una volta; dopo l'eventuale attraversamento, raggiunta una distanza massima dall'equilibrio, comincia ad avvicinarvisi lentamente senza oltrepassarlo di nuovo.

Caso 2.

$a^2 = 4bm$. Le due radici della (19) coincidono e la soluzione generale della (6) è data dalla (21). Anche qui tutte le soluzioni $x(t)$ e le loro derivate prime tendono a zero per $t \to +\infty$. Esclusa la soluzione nulla, $x(t)$ e $x'(t)$ possono annullarsi al più una volta. Il moto del punto di ascissa $x(t)$ ha lo stesso carattere del primo caso.

Caso 3.

$a^2 < 4bm$. Le radici della (19) hanno parte immaginaria non nulla. La soluzione generale è data dalla (23). Il punto oscilla lungo l'asse x con periodo costante $2\pi/\beta$, uguale per tutte le soluzioni della (6), e con ampiezza $Ce^{\alpha t}$, dove $\alpha = -a/(2m)$.

Le oscillazioni di un sistema fisico che avvengono senza l'azione di una forza esterna si chiamano oscillazioni proprie del sistema. Dalla discussione precedente risulta che, per i

sistemi degli esempi 2, 3, 4 e 5, il periodo dipende soltanto dalla struttura del sistema ed è lo stesso per tutte le oscillazioni che possono instaurarvisi. Nell'esempio 2 esso è $2\pi\sqrt{b/m - a^2/4m^2}$; nell'esempio 4 è $2\pi\sqrt{kps/v\rho l}$; nell'esempio 5 è $2\pi\sqrt{1/LC - R^2/4L^2}$.

Se $a = 0$, cioè il mezzo non oppone resistenza, l'ampiezza è costante: il punto compie oscillazioni armoniche. Se invece $a > 0$ e la resistenza è piccola, nel senso che $a^2 < 4bm$, l'ampiezza tende a zero e le oscillazioni si smorzano.

Infine, in tutti i casi la soluzione $x(t) \equiv 0$ della (6) rappresenta lo stato di quiete del punto nella posizione $x = 0$, detta posizione di equilibrio.

Se le parti reali di entrambe le radici della (19) sono negative, dalle formule (20), (21) e (23) si vede che tutte le soluzioni della (6), insieme alle loro derivate, tendono a zero per $t \to +\infty$: il moto si estingue col passare del tempo.

Se invece anche una sola radice della (19) ha parte reale positiva, esistono soluzioni della (6) che non tendono a zero per $t \to +\infty$; alcune, anzi, non rimangono nemmeno limitate. Per $m > 0$ ciò può accadere soltanto se b oppure a è negativo. Fisicamente, questo corrisponderebbe a una forza elastica che respinge il punto dall'equilibrio anziché attrarlo, oppure a una resistenza negativa del mezzo. Questi casi non si realizzano negli esempi fisici considerati all'inizio del capitolo, ma sono perfettamente realizzabili in altri modelli fisici.

Se le parti reali delle radici λ_1 e λ_2 sono nulle, il che nel caso considerato richiede $a = 0$, allora $\alpha = 0$ e, come mostra la (23), il punto compie oscillazioni armoniche con ampiezza e velocità limitate.

Equazioni lineari non omogenee a coefficienti costanti.

Consideriamo in dettaglio l'equazione

$$m\frac{d^2x}{dt^2} + a\frac{dx}{dt} + bx = A\cos\omega t. \qquad (26)$$

Essa descrive le oscillazioni lineari di un punto materiale soggetto a una forza elastica, alla resistenza del mezzo e a una forza esterna periodica $A\cos\omega t$ (si veda la (6′) del §1).

La (26) è un'equazione lineare non omogenea, mentre la (6) è l'equazione omogenea associata.

Cerchiamo la soluzione generale della (26).

Osserviamo che la somma di una soluzione dell'equazione non omogenea e di una soluzione dell'omogenea associata è ancora una soluzione della non omogenea. Per trovare la soluzione generale della (26) basta quindi determinarne una soluzione particolare qualsiasi: la soluzione generale si ottiene sommando a questa la soluzione generale dell'omogenea associata.

È naturale attendersi che il moto segua il ritmo della forza periodica esterna e cercare una soluzione particolare nella forma $x = B\cos(\omega t + \delta)$, con B e δ costanti ancora da determinare.

Scegliamole in modo che questa funzione soddisfi la (26). Calcolando $dx/dt = -B\omega \sin(\omega t + \delta)$ e $d^2x/dt^2 = -B\omega^2 \cos(\omega t + \delta)$, e sostituendo nella (26), otteniamo

$$m[-B\omega^2 \cos(\omega t + \delta)] + a[-B\omega \sin(\omega t + \delta)] + bB \cos(\omega t + \delta) = A \cos \omega t.$$

Applicando note formule trigonometriche, troviamo

$$B[(b - m\omega^2) \cos(\omega t + \delta) - a\omega \sin(\omega t + \delta)] = B\sqrt{(b - m\omega^2)^2 + a^2\omega^2} \cos(\omega t + \delta') = A \cos \omega t,$$

dove $\delta' = \delta + \gamma$ e $\gamma = \arctan[a\omega/(b - m\omega^2)]$, con la scelta del quadrante appropriato. Evidentemente, ponendo

$$\delta = -\arctan \frac{a\omega}{b - m\omega^2} \quad \text{and} \quad B = \frac{A}{\sqrt{(b - m\omega^2)^2 + a^2\omega^2}},$$

la funzione $x = B \cos(\omega t + \delta)$ soddisfa la (26).

Una soluzione della forma $B \cos(\omega t + \delta)$ esiste sempre quando $(b - m\omega^2)^2 + a^2\omega^2 \neq 0$. Se invece questa somma è nulla, cioè $a = 0$ e $b = m\omega^2$, la (26) assume la forma

$$m \frac{d^2x}{dt^2} + m\omega^2 x = A \cos \omega t.$$

Una soluzione particolare, come si verifica facilmente, è $x = (At/2\sqrt{mb}) \sin \omega t$.

Le soluzioni della non omogenea (26) descrivono oscillazioni forzate. Il fattore $\phi(\omega) = 1/\sqrt{(b - m\omega^2)^2 + a^2\omega^2}$ esprime il rapporto tra l'ampiezza B dell'oscillazione forzata e l'ampiezza A della forza eccitatrice. Il grafico di $\phi(\omega)$ si chiama curva di risonanza. La frequenza ω per cui $\phi(\omega)$ raggiunge il massimo è detta frequenza di risonanza. Calcoliamola. Se il massimo si raggiunge per $\omega_1 \neq 0$, in quel punto la derivata $\phi'(\omega)$ si annulla, cioè $-4(b - m\omega_1^2)\, m\omega_1 + 2a^2\omega_1 = 0$, da cui $\omega_1 = \sqrt{b/m - a^2/2m^2}$. Per questo valore di ω_1,

$$\phi(\omega_1) = \frac{1}{a\sqrt{b/m - a^2/4m^2}}.$$

Si vede così che, per $\omega = \omega_1$, l'ampiezza dell'oscillazione forzata cresce al diminuire di a. Per a molto piccolo, ω_1 è vicinissimo a $\sqrt{b/m}$, cioè alla frequenza delle oscillazioni libere. Se $a = 0$ e $b = m\omega^2$, abbiamo visto che l'oscillazione forzata ha la forma

$$x = \frac{At}{2\sqrt{mb}} \sin \omega t,$$

La sua ampiezza cresce dunque senza limite per $t \to +\infty$: questo è il significato matematico della risonanza. Essa si verifica quando il periodo della forza esterna coincide con quello di

un'oscillazione propria del sistema. In pratica, quando i due periodi sono vicini, gli spostamenti del sistema possono diventare molto grandi.

La possibilità di ottenere grandi oscillazioni viene spesso utilizzata nella costruzione di amplificatori, per esempio in radiotecnica. Ma grandi oscillazioni possono anche provocare la rottura di strutture, come ponti o telai di macchine. È quindi molto importante prevedere la possibilità della risonanza o di oscillazioni prossime a essa.

Come abbiamo osservato, ogni soluzione della (26) si può scrivere come somma dell'oscillazione forzata trovata e di una soluzione dell'omogenea, data dalle (20), (21) o (23). Per $a > 0$ e $b > 0$, la soluzione dell'omogenea tende a zero per $t \to +\infty$: ogni moto si avvicina quindi, col tempo, all'oscillazione forzata. Se $a = 0$ e $b > 0$, all'oscillazione forzata si sovrappone un'oscillazione propria non smorzata. Per $b = m\omega^2$ e $a = 0$ si ha risonanza.

Se sul sistema agisce una forza esterna periodica $f(t)$, le oscillazioni forzate si possono trovare così. Rappresentiamo $f(t)$, con sufficiente precisione, mediante una somma parziale di una serie trigonometrica:[2]

$$\sum_{i=1}^{n} (a_i \cos \omega_i t + b_i \sin \omega_i t). \qquad (27)$$

Determiniamo l'oscillazione forzata corrispondente a ciascun termine. Sommando queste oscillazioni otteniamo quella prodotta dalla forza $f(t)$. Se una delle frequenze presenti coincide con una frequenza propria del sistema, si ha risonanza.

§ 3. Osservazioni generali sulla formulazione e sulla risoluzione delle equazioni differenziali

Non sono molte le equazioni differenziali le cui soluzioni si possono tutte esprimere esplicitamente mediante funzioni semplici, come avviene per le equazioni lineari a coefficienti costanti. È facile dare esempi di equazioni la cui soluzione generale non si può esprimere mediante un numero finito di integrazioni di funzioni note, ossia, come si dice, mediante quadrature.

Come dimostrò Liouville nel 1841, la soluzione dell'equazione di Riccati $dy/dx + ay^2 = x^2$, con $a > 0$, non si può esprimere mediante una combinazione finita di integrali di funzioni elementari. Diventa quindi importante sviluppare metodi di approssimazione delle soluzioni applicabili ad ampie classi di equazioni.

Il fatto che in questi casi troviamo soltanto approssimazioni, anziché soluzioni esatte, non deve preoccuparci. Anzitutto possiamo calcolare tali approssimazioni, almeno in linea di principio, con qualsiasi precisione desiderata. Inoltre va sottolineato che, nella maggior parte dei casi, le stesse equazioni differenziali che descrivono un processo fisico non sono del tutto esatte, come mostrano tutti gli esempi del §1.

Un esempio particolarmente significativo è la (12), relativa al risonatore acustico. Nel ricavarla abbiamo trascurato la comprimibilità dell'aria nel collo e il movimento dell'aria nel recipiente. In realtà il moto dell'aria nel collo mette in movimento anche quella nel recipiente, ma i due moti hanno velocità e spostamenti diversi. Nel collo lo spostamento delle particelle è molto maggiore; perciò abbiamo trascurato il movimento dell'aria nel recipiente, considerandone soltanto la compressione. Per l'aria nel collo, invece, abbiamo trascurato l'energia di compressione e considerato soltanto l'energia cinetica.

Nel ricavare l'equazione del pendolo abbiamo trascurato la massa del filo. Nel ricavare la (14), relativa alle oscillazioni elettriche, abbiamo trascurato l'autoinduttanza dei collegamenti e la resistenza delle bobine. In generale, per ottenere un'equazione differenziale di un processo fisico dobbiamo sempre trascurare alcuni fattori e idealizzarne altri. Per questa ragione A. A. Andronov richiamò l'attenzione sull'importanza fisica delle equazioni le cui soluzioni cambiano poco quando le equazioni stesse subiscono piccole variazioni arbitrarie, intese in un senso opportuno. Queste equazioni, dette «strutturalmente stabili», meritano uno studio particolarmente approfondito.

Nelle indagini fisiche non sono approssimate soltanto le equazioni che descrivono le leggi di variazione delle grandezze: è approssimato persino il numero di tali grandezze. A rigore non esistono corpi rigidi. Per studiare le oscillazioni di un pendolo dovremmo quindi considerare le deformazioni del filo e del corpo che abbiamo schematizzato come un punto materiale. Analogamente, per un carico collegato a molle dovremmo considerare le masse delle singole spire. In questi esempi, però, è facile mostrare che il moto delle diverse particelle costituenti il pendolo, oppure il carico e le molle, influisce poco sul carattere delle oscillazioni. Tenerne conto renderebbe il problema tanto complesso da impedirne una soluzione con un'approssimazione adeguata; la soluzione non sarebbe allora più vicina alla realtà di quella del §1, ottenuta trascurando questi effetti. Un'idealizzazione ragionata è sempre inevitabile. Per descrivere un processo occorre considerarne gli aspetti essenziali, non tutti indistintamente: farlo complicherebbe enormemente il problema e, nella maggior parte dei casi, renderebbe impossibile calcolare una soluzione. Il compito fondamentale della fisica o della meccanica è individuare il numero minimo di grandezze che descrivono con sufficiente precisione lo stato del fenomeno in ogni istante, e formulare poi le equazioni differenziali più semplici che ne rappresentano bene le leggi di variazione. È un compito spesso molto difficile. Quali aspetti siano essenziali e quali trascurabili può essere deciso, in ultima analisi, soltanto attraverso una lunga esperienza. Solo il confronto tra le risposte dell'analisi idealizzata e i risultati sperimentali permette di giudicare la validità dell'idealizzazione.

In uno dei casi più semplici e caratteristici, il problema matematico della possibilità di ridurre il numero delle grandezze si formula così.

Supponiamo di descrivere inizialmente lo stato di un sistema fisico all'istante t mediante due grandezze $x_1(t)$ e $x_2(t)$, le cui velocità di variazione soddisfano le equazioni

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(t, x_1, x_2), \\ \epsilon \frac{dx_2}{dt} &= f_2(t, x_1, x_2), \end{aligned} \qquad (28)$$

Nella seconda equazione il coefficiente della derivata è un piccolo parametro costante ϵ. Ponendo $\epsilon = 0$, la seconda delle (28) cessa di essere un'equazione differenziale e diventa

$$f_2(t, x_1, x_2) = 0.$$

Da questa relazione ricaviamo x_2 come funzione di t e x_1, e la sostituiamo nella prima delle (28). Otteniamo l'equazione differenziale

$$\frac{dx_1}{dt} = F(t, x_1)$$

nella sola incognita x_1: il numero delle grandezze necessarie a descrivere la situazione si è ridotto a uno. Ci chiediamo ora in quali condizioni sia piccolo l'errore introdotto ponendo $\epsilon = 0$. Può infatti accadere che, per $\epsilon \to 0$, la derivata dx_2/dt cresca senza limite, cosicché il secondo membro della seconda delle (28) non tenda a zero per $\epsilon \to 0$.

§ 4. Interpretazione geometrica del problema di integrazione delle equazioni differenziali e sua generalizzazione

Per semplicità consideriamo dapprima una sola equazione differenziale del primo ordine in una funzione incognita:

$$\frac{dy}{dx} = f(x, y), \qquad (29)$$

La funzione $f(x, y)$ è definita in un dominio G del piano (x, y). In ogni punto del dominio, l'equazione assegna il coefficiente angolare della tangente al grafico di una soluzione che passa per quel punto. Se in ciascun punto (x, y) indichiamo mediante un piccolo segmento la direzione della tangente determinata da $f(x, y)$, senza distinguere i due versi, otteniamo un campo di direzioni. Il problema di trovare una soluzione della (29) con condizione iniziale $y(x_0) = y_0$ si può allora formulare così: trovare nel dominio G una curva $y = \phi(x)$ passante per $M_0(x_0, y_0)$ che abbia in ogni suo punto una tangente con il coefficiente angolare assegnato dalla (29), cioè con la direzione prescritta.

Dal punto di vista geometrico questa formulazione presenta due restrizioni poco naturali:

1. Imponendo che il coefficiente angolare della tangente in ogni punto (x, y) di G sia uguale a $f(x, y)$, escludiamo automaticamente le tangenti parallele a Oy, perché consideriamo

generalmente soltanto grandezze finite; in particolare supponiamo che la funzione al secondo membro della (29) assuma soltanto valori finiti.

2. Considerando soltanto curve che sono grafici di funzioni di x, escludiamo anche quelle che una retta perpendicolare a Ox incontra più di una volta. Infatti consideriamo soltanto funzioni a un sol valore; in particolare ogni soluzione di un'equazione differenziale è intesa come funzione univoca di x.

Generalizziamo quindi la formulazione precedente, ammettendo che in alcuni punti la tangente sia parallela a Oy. In questi punti, nei quali il coefficiente angolare rispetto a Ox non è definito, considereremo quello rispetto a Oy. In altre parole, accanto alla (29) consideriamo l'equazione

$$\frac{dx}{dy} = f_1(x, y), \qquad (29')$$

dove $f_1(x, y) = 1/f(x, y)$ quando $f(x, y) \neq 0$, e utilizziamo la seconda equazione quando la prima non ha significato. Il problema di integrare le equazioni (29) e (29′) diventa allora: trovare in G tutte le curve che hanno in ogni punto la tangente definita da queste equazioni. Le chiameremo curve integrali, o linee integrali, delle equazioni (29) e (29′), oppure del campo di tangenti da esse definito. Spesso useremo il singolare «equazione (29), (29′)» in luogo del plurale. È chiaro che il grafico di ogni soluzione della (29) è una curva integrale della (29), (29′); non vale però necessariamente il contrario. Ad esempio, una curva integrale che sia incontrata più volte da una stessa perpendicolare a Ox non è il grafico di una soluzione della (29).

Nel seguito, quando è chiaro che

$$f(x, y) = \frac{M(x, y)}{N(x, y)},$$

scriveremo soltanto l'equazione

$$\frac{dy}{dx} = \frac{M(x, y)}{N(x, y)},$$

omettendo di scrivere

$$\frac{dx}{dy} = \frac{N(x, y)}{M(x, y)}.$$

Talvolta, in luogo di queste equazioni, introdurremo un parametro t e scriveremo il sistema

$$\frac{dx}{dt} = N(x, y), \frac{dy}{dt} = M(x, y),$$

considerando x e y come funzioni di t.

Esempio 1.

L'equazione

$$\frac{dy}{dx} = \frac{y}{x} \qquad (30)$$

definisce un campo di tangenti ovunque tranne che nell'origine. Il campo è rappresentato nella figura 7. Tutte le tangenti assegnate dalla (30) passano per l'origine.

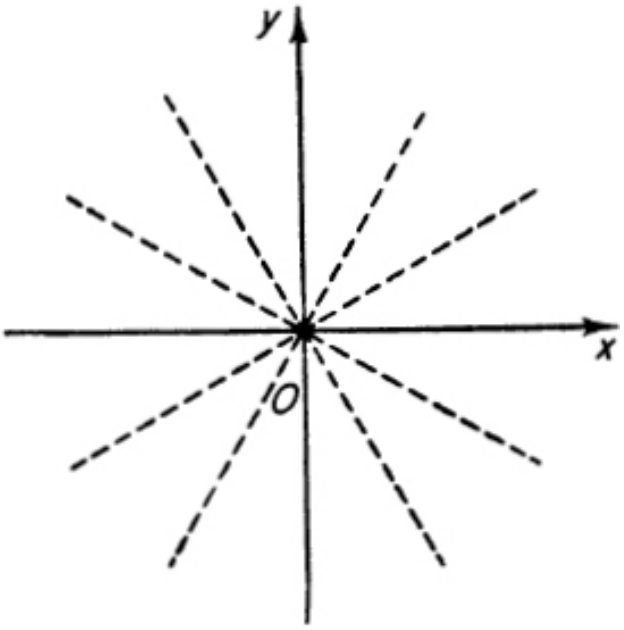


Figura 7.

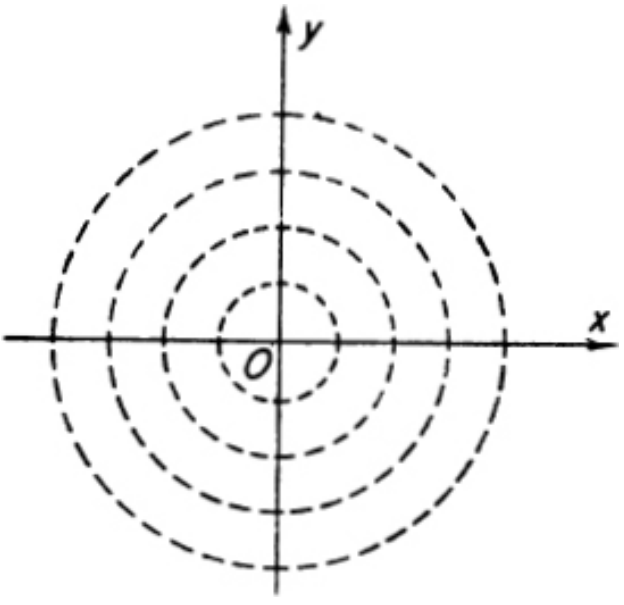


Figura 8.

È chiaro che, per ogni *k*, la funzione

$$y = kx \qquad (31)$$

è una soluzione della (30). L'insieme di tutte le curve integrali è allora definito dalla relazione

$$ax = by = 0, \qquad (32)$$

dove a e b sono costanti arbitrarie non entrambe nulle. L'asse Oy è una curva integrale della (30), ma non è il grafico di una sua soluzione.

Poiché la (30) non definisce il campo di tangenti nell'origine, le curve (31) e (32) sono, a rigore, curve integrali soltanto fuori dall'origine. È quindi più corretto dire che le curve integrali della (30) sono semirette uscenti dall'origine, anziché rette passanti per essa.

Esempio 2.

L'equazione

$$\frac{dy}{dx} = -\frac{x}{y} \qquad (33)$$

definisce un campo di tangenti ovunque tranne che nell'origine, come mostra la figura 8. In ogni punto (x, y), le tangenti definite dalle (30) e (33) sono perpendicolari. È evidente che tutte le circonferenze con centro nell'origine sono curve integrali della (33). Le soluzioni dell'equazione sono invece le funzioni

$$y = +\sqrt{R^2 - x^2},\ y = -\sqrt{R^2 - x^2},\ -R \leqslant x \leqslant R.$$

Per brevità, nel seguito diremo talvolta «una soluzione passa per il punto (x, y)» anziché, più precisamente, «il grafico di una soluzione passa per il punto (x, y)».

§ 5. Esistenza e unicità della soluzione di un'equazione differenziale; risoluzione approssimata

Il problema dell'esistenza e dell'unicità della soluzione.

Torniamo all'equazione differenziale (17), di ordine arbitrario n. In generale essa ha infinite soluzioni; per individuarne una particolare occorre aggiungere condizioni supplementari, in numero uguale all'ordine n dell'equazione. Queste condizioni possono essere molto diverse, secondo il significato fisico, meccanico o di altra natura del problema originario. Se, per esempio, dobbiamo studiare il moto di un sistema meccanico a partire da un certo stato iniziale, le condizioni si riferiscono a uno specifico valore iniziale della variabile indipendente e si chiamano condizioni iniziali. Se invece vogliamo determinare la forma di un cavo di un ponte sospeso, oppure di una trave caricata e appoggiata alle estremità, incontriamo condizioni relative a valori diversi della variabile indipendente: alle estremità del cavo o nei punti di appoggio della trave. Molti altri esempi potrebbero illustrare la varietà delle condizioni associate alle equazioni differenziali.

Supponiamo che siano assegnate le condizioni supplementari e che si debba trovare una soluzione della (17) che le soddisfi. La prima domanda è se una tale soluzione esista. Spesso non possiamo esserne certi in anticipo. Supponiamo, ad esempio, che la (17) descriva il funzionamento di un apparecchio e che vogliamo stabilire se in esso si produca un moto periodico. Le condizioni supplementari esprimeranno allora il ripetersi periodico dello stato iniziale, e non possiamo sapere a priori se esista una soluzione che le soddisfi.

Lo studio dell'esistenza e dell'unicità chiarisce, in ogni caso, quali condizioni possano essere soddisfatte per una data equazione e quali determinino la soluzione in modo univoco. Individuare queste condizioni e dimostrare l'esistenza e l'unicità della soluzione di un'equazione relativa a un problema fisico ha inoltre grande valore per la stessa teoria fisica. Mostra infatti che le ipotesi adottate nella descrizione matematica del fenomeno sono, da un lato, compatibili tra loro e, dall'altro, sufficienti a descriverlo completamente.

I metodi per studiare l'esistenza sono numerosi. Tra essi hanno particolare importanza i cosiddetti metodi diretti: si dimostra l'esistenza della soluzione costruendo soluzioni approssimate e provando che convergono alla soluzione esatta. Questi metodi non soltanto ne stabiliscono l'esistenza, ma forniscono anche un modo, anzi il principale, per approssimarla con qualunque precisione desiderata.

Nel resto della sezione considereremo, per concretezza, un problema con dati iniziali, illustrando le idee del metodo di Eulero e del metodo delle approssimazioni successive.

Il metodo delle poligonali di Eulero.

Consideriamo, in un dominio G del piano (x, y), l'equazione differenziale

$$\frac{dy}{dx} = f(x, y). \qquad (34)$$

Come abbiamo osservato, la (34) definisce in G un campo di tangenti. Scegliamo un punto (x_0, y_0) di G e tracciamo per esso la retta L_0 di coefficiente angolare $f(x_0, y_0)$. Su questa retta scegliamo un punto (x_1, y_1) abbastanza vicino al primo; nella figura 9 è indicato con 1. Tracciamo per esso la retta L_1 di coefficiente angolare $f(x_1, y_1)$ e su di essa segniamo il punto (x_2, y_2), indicato con 2. Sulla retta L_2 corrispondente a quest'ultimo punto segniamo poi (x_3, y_3), e continuiamo così, con $x_0 < x_1 < x_2 < x_3 < \cdots$. Naturalmente tutti i punti scelti devono appartenere a G. La linea spezzata che li congiunge si chiama poligonale di Eulero. La si può costruire anche nel verso delle x decrescenti; nella figura i vertici corrispondenti sono indicati con $-1, -2, -3$.

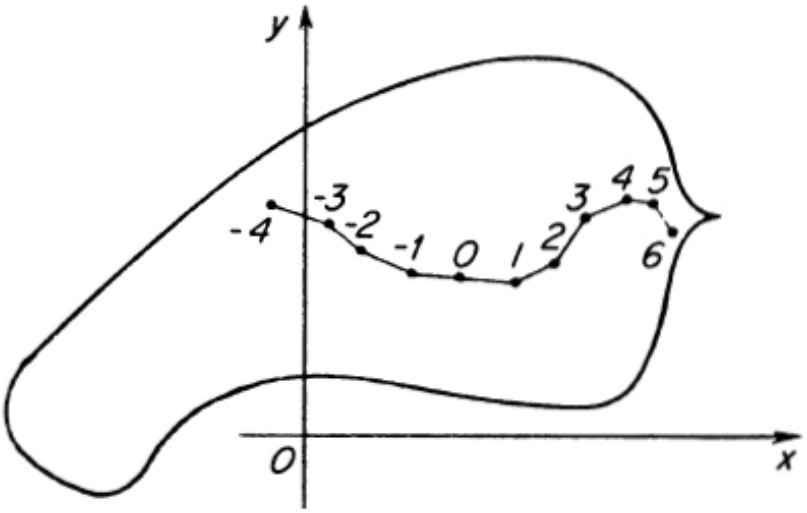


Figura 9.

È ragionevole aspettarsi che una poligonale di Eulero passante per (x_0, y_0), con segmenti abbastanza corti, rappresenti approssimativamente una curva integrale l passante per lo stesso punto; e che, accorciando i segmenti in modo che la lunghezza del più lungo tenda a zero, la poligonale si avvicini alla curva integrale.

Qui si suppone, naturalmente, che la curva integrale esista. In effetti non è difficile dimostrare che, se $f(x, y)$ è continua in G, si può trovare una successione infinita di poligonali di Eulero che converge a una curva integrale l, mentre la lunghezza dei segmenti più lunghi tende a zero. La sola continuità non garantisce però l'unicità: successioni diverse possono convergere a curve integrali diverse passanti per lo stesso punto. M. A. Lavrent'ev ha costruito un esempio di equazione della forma (29), con secondo membro continuo, nel quale la non unicità delle curve integrali si manifesta in ogni intorno di ogni punto del dominio. Perché da ciascun punto passi una sola curva integrale occorrono quindi condizioni ulteriori. È sufficiente, ad esempio, supporre che $f(x, y)$ sia continua e abbia una derivata rispetto a y limitata in tutto G. Si può allora dimostrare che per ogni punto passa una e una sola curva integrale e che ogni successione di poligonali di Eulero passanti per (x_0, y_0) converge uniformemente a questa curva, sul tratto considerato, quando la massima lunghezza dei

segmenti tende a zero. Una poligonale con segmenti sufficientemente piccoli può dunque essere assunta come approssimazione della curva integrale della (34).

Le poligonali di Eulero sostituiscono piccoli tratti delle curve integrali con segmenti delle loro tangenti. In pratica si usano spesso, al posto dei segmenti rettilinei, archi di parabole che hanno con la curva integrale un contatto di ordine superiore. In questo modo si raggiunge la stessa precisione con un numero minore di passi, cioè con meno tratti nella curva approssimante. I coefficienti dell'equazione della parabola di ordine superiore

$$y = a_0 + a_1(x - x_k) + a_2(x - x_k)^2 + \cdots + a_n(x - x_k)^n, \qquad (35)$$

che nel punto (x_k, y_k) ha un contatto di ordine n con la curva integrale della (34) passante per esso sono dati dalle formule

$$a_0 = y_k\,, \qquad (36)$$

$$a_1 = \left(\frac{dy}{dx}\right)_{x=x_k} = f(x_k\,, y_k), \qquad (36')$$

$$2a_2 = \left(\frac{d^2y}{dx^2}\right)_{x=x_k} = \left[\frac{df(x, y)}{dx}\right]_{x=x_k} = f'_x(x_k\,, y_k) + f'_y(x_k\,, y_k)\left(\frac{dy}{dx}\right)_{x=x_k}$$

$$= f'_x(x_k\,, y_k) + f'_y(x_k\,, y_k)\, f(x_k\,, y_k), \qquad (36'')$$

$$6a_3 = \left(\frac{d^3y}{dx^3}\right)_{x=x_k} = \left\{\frac{d}{dx}[f'_x(x, y(x)) + f'_y(x, y(x))\, f(x, y(x))]\right\}_{x=x_k}$$

$$= f''_{xx}(x_k\,, y_k) + 2f''_{xy}(x_k\,, y_k)\, f(x_k\,, y_k)$$

$$+ f''_{yy}(x_k\,, y_k)\, f^2(x_k\,, y_k) + f'^2_y(x_k\,, y_k)\, f(x_k\,, y_k)$$

$$+ f'_y(x_k\,, y_k)\, f'_x(x_k\,, y_k). \qquad (36''')$$

Il polinomio (35) serve soltanto a calcolarne il valore per $x = x_{k+1}$; non occorre conoscere separatamente i coefficienti a_0, a_1, a_2, ⋯, a_n. Esistono molti procedimenti per calcolare il valore del polinomio in x_{k+1}, con i coefficienti definiti dalle (36), senza determinarli esplicitamente uno per uno.

Esistono anche metodi di approssimazione delle soluzioni della (34) fondati su idee diverse. Un metodo particolarmente pratico fu sviluppato da A. N. Krylov (1863-1945).

Il metodo delle approssimazioni successive.

Descriviamo ora un altro metodo, usato quanto quello delle poligonali di Eulero. Supponiamo ancora di dover trovare una soluzione $y(x)$ della (34) che soddisfi la condizione iniziale

$$y(x_0) = y_0\,.$$

Come approssimazione iniziale scegliamo una funzione arbitraria $y_0(x)$. Per semplicità supponiamo che soddisfi anch'essa la condizione iniziale, benché non sia necessario. La sostituiamo all'incognita y nel secondo membro $f(x, y)$ e costruiamo una prima approssimazione y_1 imponendo

$$\frac{dy_1}{dx} = f[x, y_0(x)], y_1(x_0) = y_0 .$$

Poiché nel secondo membro della prima equazione compare ora una funzione nota, possiamo determinare $y_1(x)$ per integrazione:

$$y_1(x) = y_0 + \int_{x_0}^{x} f[t, y_0(t)]\, dt.$$

Ci si può aspettare che $y_1(x)$ differisca dalla soluzione cercata meno di $y_0(x)$: nel costruirla abbiamo infatti utilizzato l'equazione stessa, introducendo così una correzione nell'approssimazione iniziale. È inoltre naturale pensare che, migliorando nello stesso modo y_1, la seconda approssimazione

$$y_2(x) = y_0 + \int_{x_0}^{x} f[t, y_1(t)]\, dt$$

sia ancora più vicina alla soluzione cercata.

Supponiamo di continuare indefinitamente questo processo, costruendo la successione di approssimazioni

$$y_0(x),\ y_1(x),\ \cdots,\ y_n(x),\ \cdots .$$

Questa successione convergerà alla soluzione $y(x)$?

Un'analisi più precisa mostra che, se $f(x, y)$ è continua e f'_y è limitata in G, le funzioni $y_n(x)$ convergono effettivamente alla soluzione esatta, almeno per tutti i valori di x abbastanza vicini a x_0. Fermando il calcolo dopo un numero sufficiente di passi possiamo dunque ottenere la precisione desiderata.

Come per la singola equazione (34), possiamo approssimare anche le curve integrali di un sistema di due o più equazioni differenziali del primo ordine. La condizione essenziale è che le equazioni possano essere risolte rispetto alle derivate delle funzioni incognite. Consideriamo, ad esempio, il sistema

$$\frac{dy}{dx} = f_1(x, y, z), \frac{dz}{dx} = f_2(x, y, z). \qquad (37)$$

Supponiamo che i secondi membri siano continui e abbiano derivate limitate rispetto a y e z in un dominio G dello spazio. In queste condizioni si può dimostrare che per ogni punto (x_0, y_0, z_0) di G passa una e una sola curva integrale

$$y = \phi(x), \quad z = \psi(x)$$

del sistema (37). Le funzioni $f_1(x, y, z)$ e $f_2(x, y, z)$, insieme alla prima componente 1, determinano la direzione della tangente alla curva integrale nel punto (x, y, z). Per trovare approssimativamente $\phi(x)$ e $\psi(x)$ possiamo applicare il metodo delle poligonali di Eulero oppure altri metodi analoghi a quelli utilizzati per la (34).

Il calcolo approssimato delle soluzioni di equazioni differenziali ordinarie con condizioni iniziali può essere eseguito mediante macchine calcolatrici. Esistono calcolatori elettronici tanto rapidi che, programmati per determinare la traiettoria di un proiettile, la calcolano in un tempo minore di quello impiegato dal proiettile per raggiungere il bersaglio (si veda il capitolo XIV).

Relazione tra equazioni di ordine superiore e sistemi di equazioni del primo ordine.

Un sistema di equazioni differenziali ordinarie, risolto rispetto alla derivata di ordine più alto di ciascuna funzione incognita, si può generalmente ricondurre, introducendo nuove funzioni incognite, a un sistema del primo ordine risolto rispetto a tutte le derivate. Consideriamo, ad esempio, l'equazione

$$\frac{d^2y}{dx^2} = f\left(x, y, \frac{dy}{dx}\right). \qquad (38)$$

Poniamo

$$\frac{dy}{dx} = z. \qquad (39)$$

La (38) si può allora scrivere nella forma

$$\frac{dz}{dx} = f(x, y, z). \qquad (40)$$

A ogni soluzione della (38) corrisponde dunque una soluzione del sistema formato dalle (39) e (40). È facile verificare anche il viceversa: ogni soluzione di questo sistema determina una soluzione della (38).

Equazioni che non contengono esplicitamente la variabile indipendente.

I problemi del pendolo, del risonatore di Helmholtz, del semplice circuito elettrico e del generatore a valvola del §1 conducono a equazioni nelle quali la variabile indipendente, il tempo, non compare esplicitamente. Richiamiamo queste equazioni perché, se sono del secondo ordine, si possono ricondurre a una sola equazione del primo ordine, anziché a un sistema come nel caso generale. Questa riduzione ne semplifica molto lo studio.

Consideriamo quindi un'equazione del secondo ordine che non contenga esplicitamente l'argomento t:

$$F\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}\right) = 0. \qquad (41)$$

Poniamo

$$\frac{dx}{dt} = y \qquad (42)$$

e consideriamo y come funzione di x. Allora

$$\frac{d^2x}{dt^2} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = \frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} = y\frac{dy}{dx}.$$

e l'equazione (41) diventa

$$F\left(x, y, y\frac{dy}{dx}\right) = 0. \qquad (43)$$

In questo modo, sui tratti in cui la riduzione è applicabile, a ogni soluzione della (41) corrisponde una soluzione della (43). Viceversa, a ciascuna soluzione $y = \phi(x)$ della (43) corrispondono infinite soluzioni della (41), che si trovano integrando l'equazione

$$\frac{dx}{dt} = \phi(x), \qquad (44)$$

nella quale x è considerata come funzione di t.

È evidente che, se questa equazione è soddisfatta da $x = x(t)$, lo è anche da ogni funzione $x(t + t_0)$, con t_0 costante arbitraria.

Non tutte le curve integrali della (43) sono necessariamente grafici di una singola funzione di x: ciò accade, ad esempio, quando una curva è chiusa. Occorre allora dividerla in tratti, ciascuno dei quali sia il grafico di una funzione di x, e integrare la (44) separatamente su ciascun tratto.

I valori di x e dx/dt, che in ogni istante caratterizzano lo stato del sistema fisico corrispondente alla (41), ne definiscono la fase; il piano (x, y) si chiama quindi piano delle fasi della (41). A ogni soluzione $x = x(t)$ corrisponde la curva

$$x = x(t), \quad y = x'(t)$$

nel piano (x, y), con t considerato come parametro. Viceversa, a ogni curva integrale $y = \phi(x)$ della (43) corrisponde una famiglia di soluzioni della (41) della forma $x = x(t + t_0)$, con t_0 arbitrario. Le informazioni sul comportamento delle curve integrali della (43) si trasformano facilmente in informazioni sulle possibili soluzioni della (41). A una curva integrale chiusa della (43), per esempio, corrisponde una soluzione periodica della (41).

Applicando alla (6) la trasformazione (42), otteniamo

$$\frac{dy}{dx} = \frac{-ay - bx}{my}. \qquad (45)$$

Ponendo analogamente $v = x$ e $dv/dt = y$ nella (16), ricaviamo

$$L\frac{dy}{dx} = \frac{-[R - M(a_1 + 2a_2x + 3a_3x^2)]\,y - x}{y}. \qquad (46)$$

Come lo stato istantaneo del sistema fisico descritto dalla (41) è caratterizzato dalle due grandezze[3] x e $y = dx/dt$, così lo stato di un sistema descritto da equazioni di ordine superiore o da un sistema di equazioni differenziali è caratterizzato da un numero maggiore di grandezze. In luogo del piano delle fasi si parla allora di spazio delle fasi.

§ 6. Punti singolari

Sia P(x, y) un punto interno al dominio G nel quale consideriamo l'equazione differenziale

$$\frac{dy}{dx} = \frac{M(x, y)}{N(x, y)}. \qquad (47)$$

Se esiste un intorno R di P tale che per ogni suo punto passi una e una sola curva integrale della (47), P si chiama punto ordinario dell'equazione. Se un tale intorno non esiste, P si chiama punto singolare. Lo studio dei punti singolari è molto importante nella teoria qualitativa delle equazioni differenziali, che considereremo nella prossima sezione.

Particolare importanza hanno i punti singolari isolati, cioè quelli che possiedono un intorno privo di altri punti singolari. Nelle applicazioni li incontriamo spesso nelle equazioni della forma (47), nelle quali $M(x, y)$ e $N(x, y)$ hanno derivate continue di ordine elevato rispetto a x e y. Per tali equazioni sono ordinari tutti i punti interni al dominio nei quali almeno una delle due funzioni è diversa da zero. Consideriamo ora un punto interno (x_0, y_0) nel quale $M = N = 0$. Per semplificare la notazione supporremo $x_0 = y_0 = 0$, cosa sempre possibile traslando l'origine delle coordinate in quel punto. Sviluppando M e N mediante la formula di Taylor in potenze di x e y, e separando i termini del primo ordine, abbiamo, in un intorno di (0, 0),

$$\frac{dy}{dx} = \frac{M'_x(0, 0)\,x + M'_y(0, 0)\,y + \phi_1(x, y)}{N'_x(0, 0)\,x + N'_y(0, 0)\,y + \phi_2(x, y)}, \qquad (48)$$

dove $\phi_1(x, y)$ e $\phi_2(x, y)$ sono funzioni per le quali

$$\lim_{\substack{x\to 0\\ y\to 0}} \frac{\phi_1(x, y)}{\sqrt{x^2 + y^2}} = 0 \quad \text{and} \quad \lim_{\substack{x\to 0\\ y\to 0}} \frac{\phi_2(x, y)}{\sqrt{x^2 + y^2}} = 0.$$

Le equazioni (45) e (46) hanno questa forma. La (45) non definisce né dy/dx né dx/dy per $x = y = 0$. Se il determinante

$$\begin{vmatrix} M'_x(0,0) & M'_y(0,0) \\ N'_x(0,0) & N'_y(0,0) \end{vmatrix} \neq 0,$$

allora, qualunque valore si attribuisca a dy/dx nell'origine, questa rimane un punto di discontinuità per dy/dx e dx/dy, poiché i loro limiti dipendono dal modo in cui ci si avvicina all'origine. L'origine è dunque un punto singolare dell'equazione.

Si è dimostrato che il carattere del comportamento delle curve integrali vicino a un punto singolare isolato, qui l'origine, non è modificato dai termini ϕ_1 e ϕ_2 nel numeratore e nel denominatore, purché entrambe le radici dell'equazione

$$\begin{vmatrix} \lambda - M'_y(0,0) & -M'_x(0,0) \\ -N'_y(0,0) & \lambda - N'_x(0,0) \end{vmatrix} = 0 \qquad (49)$$

abbiano parte reale diversa da zero. Per comprendere questo comportamento studiamo quindi le curve integrali, vicino all'origine, dell'equazione

$$\frac{dy}{dx} = \frac{ax + by}{cx + dy} \qquad (50)$$

per la quale il determinante

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0.$$

La disposizione delle curve integrali nell'intorno di un punto singolare è di grande interesse in molti problemi di meccanica, per esempio nello studio delle traiettorie dei moti vicini a una posizione di equilibrio.

Si può dimostrare che è possibile scegliere in tutto il piano coordinate ξ, η, legate a x, y dalle equazioni

$$x = k_{11}\xi + k_{12}\eta,$$
$$y = k_{21}\xi + k_{22}\eta. \quad (51)$$

con coefficienti reali k_{ij}, in modo da trasformare la (50) in una delle tre forme seguenti:

$$1)\ \frac{d\eta}{d\xi} = k\frac{\eta}{\xi}, \quad \text{where} \quad k = \frac{\lambda_2}{\lambda_1}. \qquad (52)$$

$$2)\ \frac{d\eta}{d\xi} = \frac{\xi + \eta}{\xi}. \qquad (53)$$

$$3)\ \frac{d\eta}{d\xi} = \frac{\beta\xi + \alpha\eta}{\alpha\xi - \beta\eta}. \qquad (54)$$

Qui λ_1 e λ_2 sono le radici dell'equazione

$$\begin{vmatrix} c - \lambda & d \\ a & b - \lambda \end{vmatrix} = 0. \qquad (55)$$

Se le radici sono reali e distinte, la (50) si trasforma nella (52). Se coincidono, si ottiene la (52) oppure la (53), secondo che $a^2 + d^2$ sia uguale o diverso da zero. Se le radici della (55) sono complesse, $\lambda = \alpha \pm \beta i$, la trasformazione (51) porta invece alla forma (54).

Consideriamo separatamente le equazioni (52), (53) e (54), premettendo alcune osservazioni.

Anche se gli assi Ox e Oy sono perpendicolari, gli assi Oξ e Oη non lo sono necessariamente. Per semplificare le figure li rappresenteremo tuttavia come perpendicolari. Inoltre la trasformazione (51) può modificare le scale sui due assi; anche in questo caso, per semplicità, supporremo che esse rimangano invariate. Così, ad esempio, al posto delle circonferenze concentriche della figura 8 potrebbe comparire una famiglia di ellissi simili, con la stessa orientazione e con centro comune nell'origine.

Tutte le curve integrali della (52) sono date da una relazione della forma

$$a\eta + b\,|\,\xi\,|^k = 0,$$

dove a e b sono costanti arbitrarie.

La figura 10 rappresenta le curve integrali della (52) nel caso $k > 1$. Tutte, tranne l'asse Oη, sono tangenti nell'origine all'asse Oξ. Il caso $0 < k < 1$ si riconduce al precedente scambiando ξ e η, cioè i ruoli dei due assi. Per $k = 1$ la (52) diventa la (30), le cui curve integrali sono rappresentate nella figura 7.

La figura 11 mostra le curve integrali della (52) per $k < 0$. In questo caso soltanto le curve situate sugli assi Oξ e Oη raggiungono il punto O. Tutte le altre, dopo essersi avvicinate all'origine fino a una certa distanza minima, se ne allontanano di nuovo.

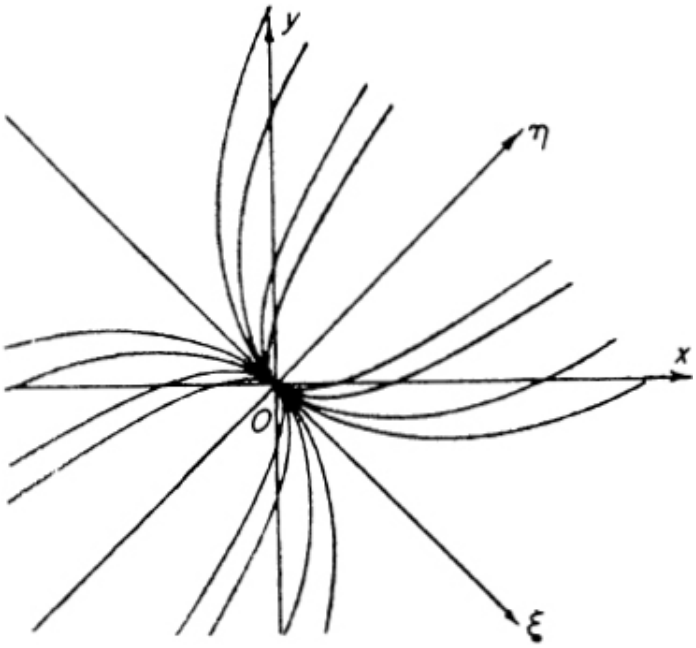


Figura 10.

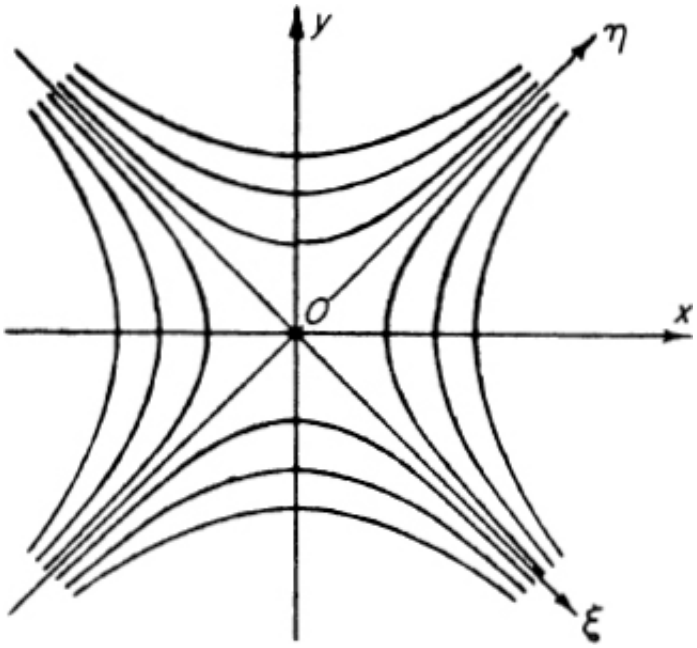


Figura 11.

In questo caso O si chiama punto di sella, perché le curve integrali ricordano le linee di livello di una carta topografica in corrispondenza di un valico montano, o sella.

Tutte le curve integrali della (53) sono date dall'equazione

$$b\eta = \xi(a + b \ln | \xi |),$$

dove a e b sono costanti arbitrarie. Sono rappresentate schematicamente nella figura 12 e sono tutte tangenti nell'origine all'asse Oη.

Se ogni curva integrale che entra in un certo intorno del punto singolare O raggiunge questo punto con una direzione determinata, cioè con una tangente definita nell'origine, come nelle figure 10 e 12, O si chiama nodo.

La (54) si integra più facilmente passando alle coordinate polari ρ e ϕ, mediante le sostituzioni

$$\xi = \rho \cos \phi, \quad \eta = \rho \sin \phi.$$

L'equazione diventa allora

$$\frac{d\rho}{d\phi} = k\rho, \quad \text{where} \quad k = \frac{\alpha}{\beta},$$

e quindi

$$\rho = Ce^{k\phi}. \tag{56}$$

Se $k > 0$, tutte le curve integrali si avvicinano a O avvolgendosi infinite volte attorno a esso per $\phi \to -\infty$ (figura 13). Se $k < 0$, questo avviene invece per $\phi \to +\infty$.

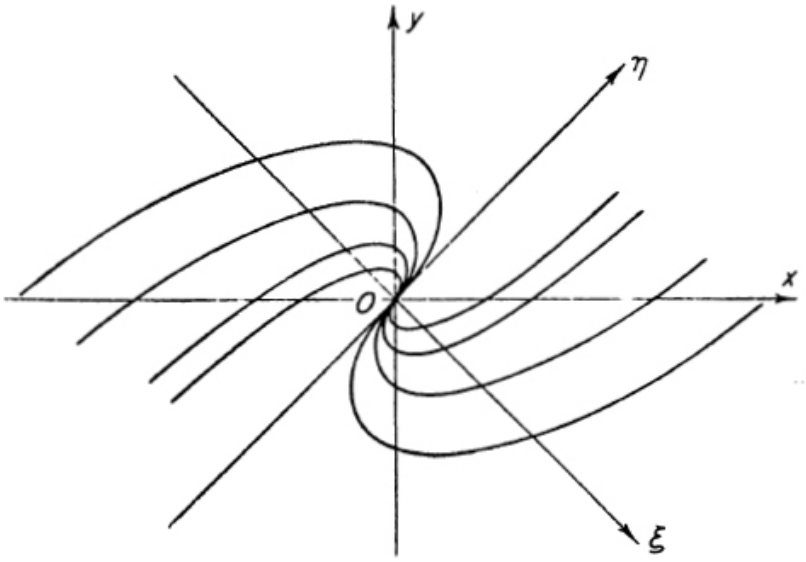


Figura 12.

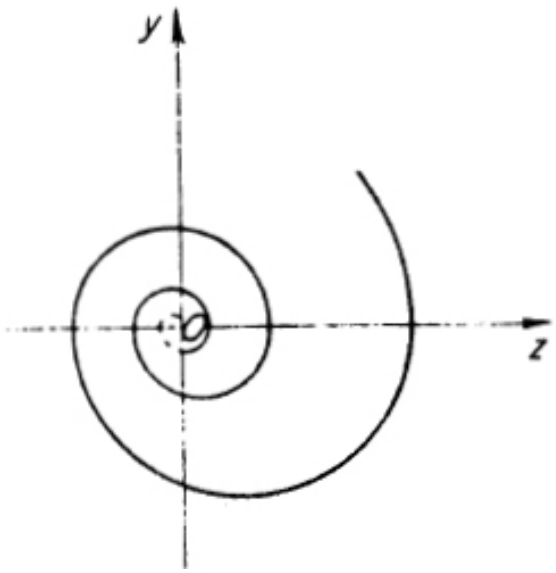


Figura 13.

In entrambi i casi O si chiama fuoco. Se invece $k = 0$, le curve integrali della (56) sono circonferenze con centro in O. In generale, un punto O si chiama centro se un suo intorno è interamente riempito da curve integrali chiuse che lo circondano.

Un centro si può facilmente trasformare in un fuoco aggiungendo al numeratore e al denominatore del secondo membro della (54) termini di ordine arbitrariamente elevato. In questo caso, quindi, i termini del primo ordine non determinano il comportamento delle curve integrali vicino al punto singolare.

L'equazione (55) associata alla (45) coincide con l'equazione caratteristica (19). Le figure 10 e 12 rappresentano pertanto, schematicamente, nel piano delle fasi (x, y), il comportamento delle curve

$$x = x(t), \quad y = x'(t),$$

corrispondenti alle soluzioni della (6) quando λ_1 e λ_2 sono reali e dello stesso segno. La figura 11 corrisponde a radici reali di segno opposto, mentre le figure 13 e 8, quest'ultima nel caso del centro, corrispondono a radici complesse. Se le parti reali di entrambe le radici sono negative, il punto $(x(t), y(t))$ tende all'origine per $t \to +\infty$: lo stato $x = y = 0$ è allora un equilibrio stabile. Se invece almeno una delle radici ha parte reale positiva, l'equilibrio nell'origine non è stabile.

§ 7. Teoria qualitativa delle equazioni differenziali ordinarie

Una parte importante della teoria generale delle equazioni differenziali ordinarie è la teoria qualitativa, nata alla fine del secolo scorso dalle esigenze della meccanica e dell'astronomia.

In molti problemi pratici occorre stabilire il carattere delle soluzioni di un'equazione che descrive un processo fisico e studiarne le proprietà mentre la variabile indipendente percorre un intervallo finito o infinito. Nella meccanica celeste, che studia il moto dei corpi celesti, è importante, ad esempio, conoscere il comportamento delle soluzioni delle equazioni del moto dei pianeti e degli altri corpi per intervalli di tempo illimitati.

Come abbiamo detto, soltanto per poche equazioni particolarmente semplici la soluzione generale si può esprimere mediante integrali di funzioni note. È sorto quindi il problema di studiare le proprietà delle soluzioni direttamente dall'equazione. Poiché una soluzione è rappresentata da una curva nel piano o nello spazio, si tratta di indagare le proprietà delle curve integrali, la loro disposizione e il loro comportamento vicino ai punti singolari. Rimangono in una parte limitata del piano oppure hanno rami che vanno all'infinito? Alcune sono chiuse? Lo studio di queste e altre domande costituisce la teoria qualitativa delle equazioni differenziali.

I fondatori di questa teoria sono il matematico russo A. M. Ljapunov e il matematico francese H. Poincaré.

Nella sezione precedente abbiamo esaminato in dettaglio uno dei suoi problemi importanti: la disposizione delle curve integrali nell'intorno di un punto singolare. Passiamo ora ad altri problemi fondamentali.

Stabilità.

Negli esempi iniziali del capitolo, la stabilità o instabilità dell'equilibrio si riconosceva facilmente mediante considerazioni fisiche, senza studiare le equazioni. Nell'esempio 3 è evidente che, spostando il pendolo dalla posizione di equilibrio OA a una posizione vicina OA′ mediante una forza esterna, cioè modificando di poco le condizioni iniziali, il moto successivo non può portarlo molto lontano dall'equilibrio. Lo scostamento sarà tanto minore quanto più piccolo è quello iniziale: la posizione di equilibrio è quindi stabile.

In casi più complessi la questione è molto più difficile e può essere affrontata soltanto studiando le equazioni differenziali corrispondenti. La stabilità dell'equilibrio è strettamente legata alla stabilità del moto. Risultati fondamentali in questo campo furono ottenuti da A. M. Ljapunov.

Supponiamo che un processo fisico sia descritto dal sistema

$$\frac{dx}{dt} = f_1(x, y, t),$$
$$\frac{dy}{dt} = f_2(x, y, t). \qquad (57)$$

Per semplicità consideriamo soltanto due equazioni, ma le conclusioni restano valide anche per sistemi più grandi. Seguendo l'uso di Ljapunov, chiameremo talvolta moto ogni soluzione particolare della (57), costituita dalle due funzioni $x(t)$ e $y(t)$. Supporremo che $f_1(x, y, t)$ e $f_2(x, y, t)$ abbiano derivate parziali continue. In questo caso, come si è dimostrato, la soluzione è univocamente determinata assegnando, a un istante $t = t_0$, i valori iniziali $x(t_0) = x_0$ e $y(t_0) = y_0$.

Indichiamo con $x(t, x_0, y_0)$ e $y(t, x_0, y_0)$ la soluzione del sistema (57) che soddisfa le condizioni iniziali

$$x = x_0 \text{ and } y = y_0 \text{ for } t = t_0 .$$

La soluzione si chiama stabile nel senso di Ljapunov se, per ogni $t > t_0$, le funzioni $x(t, x_0, y_0)$ e $y(t, x_0, y_0)$ subiscono variazioni arbitrariamente piccole quando si modificano abbastanza poco i valori iniziali x_0 e y_0.

Più precisamente, affinché la soluzione sia stabile nel senso di Ljapunov, le differenze

$$| x(t, x_0 + \delta_1 , y_0 + \delta_2) - x(t, x_0 , y_0) |,$$
$$| y(t, x_0 + \delta_1 , y_0 + \delta_2) - y(t, x_0 , y_0) | \qquad (58)$$

devono essere minori, in valore assoluto, di un qualunque numero positivo prefissato ε, per tutti i tempi $t > t_0$, purché δ_1 e δ_2 siano abbastanza piccoli in valore assoluto.

Ogni moto che non è stabile nel senso di Ljapunov si chiama instabile.

Ljapunov chiamava imperturbato il moto $x(t, x_0, y_0)$, $y(t, x_0, y_0)$ e perturbato il moto $x(t, x_0 + \delta_1, y_0 + \delta_2)$, $y(t, x_0 + \delta_1, y_0 + \delta_2)$, con condizioni iniziali vicine. La stabilità significa dunque che il moto perturbato deve differire poco da quello imperturbato per tutti i tempi successivi a t_0.

La stabilità dell'equilibrio è un caso particolare della stabilità del moto, nel quale il moto imperturbato è

$$x(t, x_0 , y_0) \equiv 0 \text{ and } y(t, x_0 , y_0) \equiv 0.$$

Viceversa, la stabilità di un qualunque moto $x = \phi_1(t)$, $y = \phi_2(t)$ del sistema (57) si può ricondurre alla stabilità dell'equilibrio di un altro sistema. Basta sostituire alle incognite $x(t)$ e $y(t)$ le nuove funzioni

$$\xi = x - \phi_1(t) \text{ and } \eta = y - \phi_2(t). \qquad (59)$$

Nel sistema così trasformato il moto $x = \phi_1(t)$, $y = \phi_2(t)$ corrisponde a $\xi = \eta = 0$, cioè a una posizione di equilibrio. Nel seguito supporremo sempre eseguita la trasformazione (59), così da studiare la stabilità nel senso di Ljapunov soltanto per la soluzione $x = y = 0$.

La condizione di stabilità significa ora che, scegliendo δ_1 e δ_2 abbastanza piccoli, la traiettoria del moto perturbato nel piano (x, y) rimane, per ogni $t > t_0$, dentro il quadrato di lato 2ε, con lati paralleli agli assi e centro nell'origine, per quanto piccolo sia $\varepsilon > 0$.

Ci interessano i casi nei quali possiamo trarre conclusioni sulla stabilità o instabilità senza conoscere un integrale del sistema. La stabilità ha grande importanza pratica nel moto dei proiettili e degli aeroplani; quella delle orbite è importante nella meccanica celeste, dove il moto dei pianeti e degli altri corpi conduce proprio a questo tipo di indagine.

Supponiamo che $f_1(x, y, t)$ e $f_2(x, y, t)$ si possano rappresentare nella forma

$$\begin{aligned} f_1(x, y, t) &= a_{11}x + a_{12}y + R_1(x, y, t), \\ f_2(x, y, t) &= a_{21}x + a_{22}y + R_2(x, y, t), \end{aligned} \tag{60}$$

dove gli a_{ij} sono costanti e $R_1(x, y, t)$, $R_2(x, y, t)$ sono funzioni tali che

$$| R_1(x, y, t) | \leqslant M(x^2 + y^2) \text{ and } | R_2(x, y, t) | \leqslant M(x^2 + y^2), \tag{61}$$

con M costante positiva.

Sostituendo le (60) nel sistema (57) e trascurando R_1 e R_2, otteniamo il sistema a coefficienti costanti

$$\begin{aligned} \frac{dx}{dt} &= a_{11}x + a_{12}y, \\ \frac{dy}{dt} &= a_{21}x + a_{22}y, \end{aligned} \tag{62}$$

detto sistema di prima approssimazione del sistema non lineare (57).

Prima di Ljapunov, i ricercatori si limitavano a studiare la stabilità della prima approssimazione, ritenendo che le conclusioni valessero anche per il sistema non lineare originario. Ljapunov mostrò per primo che questa conclusione è falsa in generale. Diede però condizioni molto ampie nelle quali la prima approssimazione risolve completamente il problema della stabilità del sistema non lineare. Una di esse è la seguente: se entrambe le radici dell'equazione

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0$$

hanno parte reale negativa e le funzioni R_1 e R_2 soddisfano la (61), la soluzione $x(t) \equiv 0$, $y(t) \equiv 0$ è stabile nel senso di Ljapunov. Se almeno una radice ha parte reale positiva, questa

soluzione è invece instabile, sempre sotto le condizioni (61). Ljapunov fornì anche numerose altre condizioni sufficienti per la stabilità e per l'instabilità di un moto.[4]

Se i secondi membri della (57) non dipendono da *t*, dividendo la prima equazione per la seconda otteniamo

$$\frac{dy}{dx} = \frac{f_1(x, y)}{f_2(x, y)}. \qquad (63)$$

L'origine è un punto singolare di questa equazione. Nel caso di equilibrio stabile può essere un fuoco, un nodo o un centro, ma non una sella.

La stabilità o instabilità dell'equilibrio fornisce dunque informazioni sul carattere del punto singolare.

Il comportamento globale delle curve integrali.

Talvolta è importante costruire una rappresentazione schematica del comportamento delle curve integrali «in grande», cioè nell'intero dominio del sistema, senza preoccuparsi di conservarne la scala. Consideriamo lo spazio nel quale il sistema definisce un campo di direzioni come lo spazio delle fasi di un processo fisico. Lo schema complessivo delle curve integrali ci darà allora un'idea di tutti i processi, o moti, possibili nel sistema. Nelle figure 10-13 abbiamo già costruito rappresentazioni schematiche del comportamento delle curve vicino a un punto singolare isolato.

Uno dei problemi fondamentali della teoria consiste nel trovare un metodo il più semplice possibile per costruire uno schema del comportamento dell'intera famiglia di curve integrali in tutto il dominio di definizione: studiarle, cioè, «in grande». Questo problema è ancora quasi inesplorato negli spazi di dimensione superiore a due ed è ben lontano dall'essere risolto persino per una sola equazione della forma

$$\frac{dy}{dx} = \frac{M(x, y)}{N(x, y)} \qquad (64)$$

anche quando $M(x, y)$ e $N(x, y)$ sono polinomi.

Nel seguito supporremo che M e N abbiano derivate parziali prime continue.

Se tutti i punti di un dominio semplicemente connesso G, nel quale è definito il secondo membro della (64), sono ordinari, la famiglia delle curve integrali si può rappresentare schematicamente come una famiglia di segmenti di rette parallele: per ogni punto passa infatti una curva integrale e due curve integrali non possono intersecarsi. Per un'equazione più generale, con punti singolari, la struttura può essere molto più complessa. Il caso di infiniti punti singolari, nei quali numeratore e denominatore si annullano insieme, può essere eliminato almeno quando M e N sono polinomi, rimuovendo gli eventuali fattori comuni. Limitiamoci dunque al caso di un numero finito di punti singolari isolati. Il comportamento delle curve vicino a questi punti è un elemento essenziale nella costruzione di uno schema complessivo.

Un altro elemento tipico di tale schema è costituito dai cosiddetti cicli limite. Consideriamo l'equazione

$$\frac{d\rho}{d\phi} = \rho - 1, \qquad (65)$$

dove ρ e ϕ sono coordinate polari nel piano (x, y).

L'insieme delle curve integrali della (65) è dato dalla formula

$$\rho = 1 + Ce^{\phi}, \qquad (66)$$

dove C è una costante arbitraria, diversa per le diverse curve integrali. Perché ρ sia non negativo, quando $C < 0$ occorre che ϕ non superi $-\ln|C|$. La famiglia comprende:

1. la circonferenza $\rho = 1$, per $C = 0$;

2. le spirali uscenti dall'origine che si avvicinano alla circonferenza dall'interno per $\phi \to -\infty$, con $C < 0$;

3. le spirali che si avvicinano alla circonferenza $\rho = 1$ dall'esterno per $\phi \to -\infty$, con $C > 0$ (figura 14).

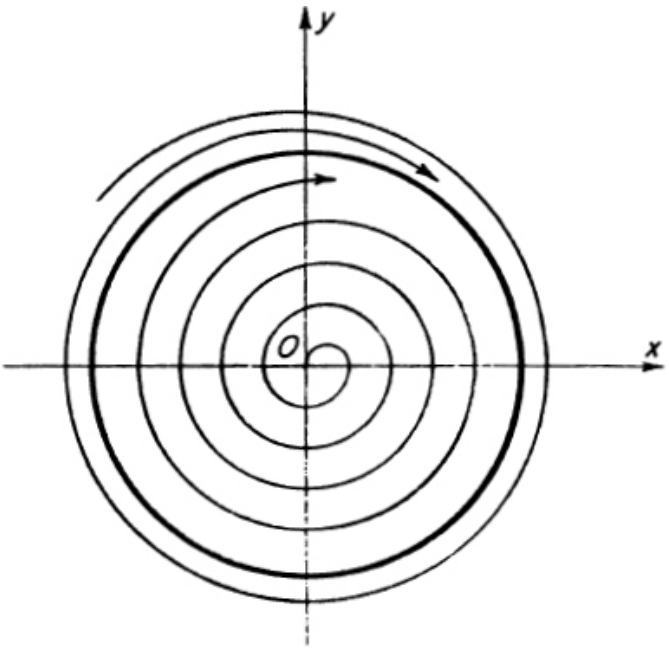


Figura 14.

La circonferenza $\rho = 1$ si chiama ciclo limite della (65). In generale, una curva integrale chiusa l è un ciclo limite se possiede un intorno anulare privo di punti singolari e interamente riempito, eccettuata l stessa, da curve integrali non chiuse.

Dalla (65) si vede che tutti i punti della circonferenza sono ordinari: un piccolo tratto di un ciclo limite non si distingue quindi da un piccolo tratto di qualunque altra curva integrale.

Ogni curva integrale chiusa nel piano (x, y) fornisce una soluzione periodica $[x(t), y(t)]$ del sistema

$$\frac{dx}{dt} = N(x, y), \frac{dy}{dt} = M(x, y), \qquad (67)$$

che descrive la legge di evoluzione di un sistema fisico. Le curve integrali del piano delle fasi che, per $t \to +\infty$, si avvicinano a un ciclo limite rappresentano moti che si avvicinano a moti periodici.

Supponiamo che per ogni punto (x_0, y_0) sufficientemente vicino a un ciclo limite l avvenga quanto segue: scegliendolo come punto iniziale della soluzione del sistema (67), la curva descritta da $[x(t), y(t)]$ si avvicina a l per $t \to +\infty$. Il moto diventa cioè approssimativamente periodico. In questo caso il ciclo limite si chiama stabile. Le oscillazioni che si comportano così corrispondono fisicamente alle autooscillazioni. In alcuni sistemi autooscillanti possono esistere più regimi oscillatori stabili, di ampiezza diversa: le condizioni iniziali determinano quale si instauri. Se tali sistemi sono descritti da equazioni della forma (67), nel piano delle fasi esistono i corrispondenti cicli limite.

Il problema di trovare, anche solo approssimativamente, i cicli limite di un'equazione differenziale non è ancora risolto in modo soddisfacente. Il metodo più usato è quello proposto da Poincaré, basato sulla costruzione di «cicli senza contatto» e sul teorema seguente. Supponiamo di trovare nel piano due curve chiuse L_1 e L_2 con queste proprietà:

1. L_2 è contenuta nella regione delimitata da L_1.

2. Nella regione anulare Ω compresa tra le due curve non vi sono punti singolari della (64).

3. Entrambe le curve hanno tangente in ogni punto e la direzione della tangente non coincide mai con quella assegnata dal campo della (64).

4. In tutti i punti di entrambe le curve, il coseno dell'angolo tra la normale rivolta all'interno di Ω e il vettore di componenti $[N(x, y), M(x, y)]$ mantiene uno stesso segno.

Allora tra L_1 e L_2 esiste almeno un ciclo limite della (64).

Poincaré chiamò L_1 e L_2 cicli senza contatto.

La dimostrazione si fonda sul seguente fatto, piuttosto intuitivo. Supponiamo che, al diminuire di t, oppure al suo aumentare, tutte le curve integrali

$$x = x(t), \quad y = y(t)$$

della (64), o equivalentemente del sistema (67) con parametro t, che incontrano L_1 o L_2 entrino nella regione anulare Ω. Esse devono allora tendere a una curva chiusa l situata tra le due frontiere, perché nessuna curva integrale interna all'anello può uscirne e non vi sono punti singolari.

Anche trovare cicli senza contatto è però difficile e non se ne conoscono metodi generali. In casi particolari è stato possibile costruirli e dimostrare così l'esistenza dei cicli limite.

In radiotecnica è importante trovare i cicli limite, cioè i regimi autooscillatori, dell'equazione (16) del generatore a valvola. Per equazioni di questo tipo N. M. Krylov e N. N. Bogoljubov elaborarono, circa vent'anni fa, un metodo per calcolare approssimativamente un ciclo limite esistente. All'incirca nello stesso periodo, i fisici sovietici L. I. Mandel'štam, N. D. Papaleksi e A. A. Andronov giustificarono l'applicabilità del cosiddetto metodo del piccolo parametro, già in parte impiegato nella pratica ma senza una dimostrazione rigorosa. Andronov fu inoltre il primo ad applicare sistematicamente all'analisi dei sistemi autooscillanti i metodi teorici di Ljapunov e Poincaré, ottenendo numerosi risultati importanti.

Come abbiamo ricordato nel §3, i sistemi strutturalmente stabili hanno un ruolo importante nella fisica. Andronov e L. S. Pontrjagin compilarono un catalogo degli elementi con i quali costruire un quadro completo delle curve integrali, nel piano, di un'equazione strutturalmente stabile della forma (64). Era noto da tempo, ad esempio, che un centro può essere distrutto da piccole variazioni dell'equazione. Nello schema di un'equazione strutturalmente stabile non può quindi comparire un centro, cioè una famiglia di curve integrali chiuse che circondano un punto singolare.

Il problema del comportamento globale delle curve integrali è ancora lontano da una soluzione definitiva. Osserviamo che anche il problema analogo, e probabilmente più semplice, della forma delle curve algebriche reali del piano, definite da

$$P(x, y) = 0,$$

con $P(x, y)$ polinomio di grado n, è lontano da una soluzione completa. La forma di queste curve è nota completamente soltanto per $n < 6$.

Le soluzioni del sistema associato alla (64) definiscono moti nel piano. Sostituendo a ogni punto (x_0, y_0) il punto corrispondente $[x(t, x_0, y_0), y(t, x_0, y_0)]$, dove le due funzioni sono la soluzione con condizioni iniziali $x = x_0$, $y = y_0$ all'istante t_0, otteniamo una trasformazione del piano dipendente dal parametro t. Trasformazioni analoghe, e i moti da esse generati, si possono considerare su una sfera, su un toro o su altre varietà. Le proprietà di questi moti sono oggetto della teoria dei sistemi dinamici. Nell'intorno di ogni punto essi sono soluzioni di un sistema di equazioni differenziali. Nell'ultimo decennio questa teoria si è sviluppata su vasta scala nelle opere di V. V. Stepanov, A. Ja. Hinčin, N. N. Bogoljubov, N. M. Krylov, A. A. Markov, V. V. Nemyckij e altri, nonché nei lavori di G. D. Birkhoff e di altri matematici.

In questo capitolo abbiamo delineato brevemente lo stato attuale della teoria delle equazioni differenziali ordinarie e cercato di descriverne i problemi. L'esposizione non pretende in alcun modo di essere completa. Abbiamo dovuto omettere molti rami della teoria che nascono dallo studio di problemi più speciali o richiedono conoscenze matematiche più ampie di quelle presupposte nel lettore. Non abbiamo, ad esempio, affrontato l'importante teoria delle equazioni differenziali con argomento complesso. Non abbiamo potuto esaminare nemmeno la teoria dei problemi al contorno e, in particolare, delle autofunzioni, di grande importanza nelle applicazioni.

Abbiamo inoltre dedicato poco spazio ai metodi approssimati di risoluzione numerica o analitica delle equazioni differenziali. Per questi argomenti rimandiamo il lettore alla letteratura specializzata.

Letture consigliate

R. P. Agnew, Differential equations, 2nd ed., McGraw-Hill, New York, 1960.

E. A. Coddington, An introduction to ordinary differential equations, Prentice-Hall, Englewood Cliffs, N. J., 1961.

E. A. Coddington and N. Levinson, Theory of ordinary differential equations, McGraw-Hill, New York, 1955.

W. Hurewicz, Lectures on ordinary differential equations, Technology Press and Wiley, New York, 1958.

S. Lefschetz, Differential equations: geometric theory, Interscience, New York, 1957.

Note al capitolo V

[1] La somma delle radici λ_1 e λ_2 della quadratica (19) è $\lambda_1 + \lambda_2 = -a/m$. Se le radici coincidono, vale quindi la seconda delle relazioni precedenti.

[2] Si veda il capitolo XII, §7.

[3] I valori di d^2x/dt^2, d^3x/dt^3, ⋯ nello stesso istante sono determinati da x e dx/dt, attraverso la (41) e le equazioni ottenute derivandola (si confronti la formula (36)).

[4] A. M. Ljapunov, *The general problem of stability of motion* («Il problema generale della stabilità del moto»).

Nota editoriale

La traduzione conserva la struttura, le quattordici figure e le formule del capitolo. Sono corretti errori di trascrizione negli indici e nelle lettere greche, la scrittura della derivata della funzione inversa nel §1 e alcuni rinvii interni. Nella trasformazione lineare (51), ricomposta tipograficamente, il coefficiente di ξ nella seconda riga è indicato con k_{21}, anziché ripetere k_{12}. Nella definizione di stabilità il lato del quadrato è 2ε; nella definizione di ciclo limite si intende un intorno anulare della curva, non un disco interamente privo di punti singolari.

Sono rese esplicite alcune precisazioni: il decadimento dipende dalla quantità di radio presente in ogni istante; nel pendolo il vincolo impedisce lo spostamento radiale, senza che la componente radiale del peso debba essere interamente equilibrata durante il moto; nella prova di unicità del §2 il ragionamento energetico si applica direttamente per $t \geq t_0$, mentre per i tempi precedenti si usa l'inversione del tempo e la trasformazione esponenziale indicata nel testo. Sono escluse le soluzioni identicamente nulle dalle affermazioni sul numero degli zeri. Per la fase della risposta forzata si sceglie il quadrante appropriato dell'arcotangente.

Nell'esempio del generatore a valvola, la relazione $dv/dt = I$ e la formula (16) sono conservate nella normalizzazione dell'originale: la capacità è implicitamente posta uguale a uno, oppure assorbita nei coefficienti. Con capacità C esplicita si usa $I = C\, dv/dt$. L'espressione «strutturalmente stabile» rende il termine inglese «insensitive» riferito ai sistemi di Andronov. Le affermazioni sullo stato delle ricerche e i riferimenti temporali rispecchiano l'epoca dell'opera.

CAPITOLO VI

EQUAZIONI ALLE DERIVATE PARZIALI

S. L. Sobolev e O. A. Ladyženskaja

§1. Introduzione

Nello studio dei fenomeni naturali si incontrano equazioni alle derivate parziali con la stessa frequenza delle equazioni differenziali ordinarie. Ciò avviene, di regola, quando un fenomeno è descritto da una funzione di più variabili. Dallo studio della natura è nata la classe di equazioni alle derivate parziali che oggi è la più approfonditamente studiata e probabilmente la più importante nell'insieme delle conoscenze umane: le equazioni della fisica matematica.

Consideriamo anzitutto le oscillazioni in un mezzo qualsiasi. Durante tali oscillazioni, ogni punto del mezzo che in equilibrio occupa la posizione (x, y, z) subisce all'istante t uno spostamento rappresentato da un vettore $u(x, y, z, t)$, dipendente dalla posizione iniziale del punto e dal tempo t. Il processo è dunque descritto da un campo vettoriale. È però facile vedere che la conoscenza di questo campo, cioè del campo degli spostamenti dei punti del mezzo, non basta a descrivere completamente l'oscillazione. Occorre conoscere anche, per esempio, la densità $\rho(x, y, z, t)$ in ciascun punto, la temperatura $T(x, y, z, t)$ e gli sforzi interni, ossia le forze che tutta la parte restante del corpo esercita su un suo volume scelto arbitrariamente.

Gli eventi e i processi fisici che si svolgono nello spazio e nel tempo consistono sempre nelle variazioni temporali di determinate grandezze fisiche associate ai punti dello spazio. Come abbiamo visto nel capitolo II, queste grandezze si possono descrivere mediante funzioni di quattro variabili indipendenti, x, y, z e t, dove le prime tre sono le coordinate di un punto dello spazio e t è il tempo.

Le grandezze fisiche possono essere di natura diversa. Alcune sono completamente caratterizzate dal loro valore numerico, come la temperatura, la densità e così via, e si chiamano scalari. Altre possiedono una direzione e sono quindi grandezze vettoriali: velocità, accelerazione, intensità del campo elettrico, ecc. Una grandezza vettoriale si può esprimere non soltanto mediante la lunghezza del vettore e la sua direzione, ma anche mediante le sue «componenti», scomponendola nella somma di tre vettori mutuamente perpendicolari, per esempio paralleli agli assi coordinati.

Nella fisica matematica una grandezza scalare, o un campo scalare, è rappresentata da una funzione di quattro variabili indipendenti; una grandezza vettoriale definita in tutto lo spazio, ossia un campo vettoriale, è invece descritta da tre funzioni di queste variabili. Possiamo scrivere una tale grandezza nella forma

$$\boldsymbol{u}(x, y, z, t),$$

dove il carattere neretto indica che u è un vettore, oppure mediante tre funzioni

$$u_x(x, y, z, t), \quad u_y(x, y, z, t), \quad u_z(x, y, z, t),$$

dove u_x, u_y e u_z denotano le proiezioni del vettore sugli assi coordinati.

Oltre alle grandezze vettoriali e scalari, in fisica compaiono enti ancora più complessi, come lo stato di sforzo di un corpo in un dato punto. Queste grandezze si chiamano tensori; fissati gli assi coordinati, esse possono essere caratterizzate in ogni punto da un insieme di funzioni delle medesime quattro variabili indipendenti.

La descrizione di fenomeni fisici molto diversi viene dunque fornita, di solito, mediante più funzioni di più variabili. Naturalmente una simile descrizione non può essere assolutamente esatta.

Per esempio, descrivendo la densità di un mezzo mediante una funzione delle nostre variabili indipendenti, trascuriamo il fatto che in un dato punto non si può attribuire alla densità un valore qualsiasi. I corpi che studiamo hanno struttura molecolare e le molecole non sono contigue, ma si trovano a distanze finite le une dalle altre. Per lo più queste distanze sono sensibilmente maggiori delle dimensioni delle molecole stesse. La densità considerata è quindi il rapporto fra la massa contenuta in un volume piccolo, ma non estremamente piccolo, e questo stesso volume. La densità in un punto viene abitualmente concepita come il limite di tali rapporti al diminuire del volume. Una semplificazione e un'idealizzazione ancora maggiori intervengono nel concetto di temperatura di un mezzo. Il calore di un corpo è legato al moto disordinato delle sue molecole. Le energie delle singole molecole sono diverse; ma, considerando un volume che ne contenga un gran numero, l'energia media dei loro moti disordinati definisce ciò che si chiama temperatura.

Analogamente, parlando della pressione di un gas o di un liquido sulla parete di un recipiente, non dobbiamo immaginare che una particella prema effettivamente e stabilmente contro la parete. Nel loro moto disordinato le particelle urtano la parete e rimbalzano. Ciò che descriviamo come pressione è quindi costituito da un grandissimo numero di impulsi ricevuti da una porzione di parete piccola dal punto di vista ordinario, ma enorme rispetto alle distanze fra le molecole del liquido o del gas. Si potrebbero facilmente dare decine di esempi analoghi. La maggior parte delle grandezze studiate in fisica ha proprio questo carattere. La fisica matematica tratta grandezze idealizzate, astraendo dalle proprietà concrete degli enti fisici corrispondenti e considerando soltanto i valori medi di queste grandezze.

Una simile idealizzazione può sembrare piuttosto grossolana, ma, come vedremo, è molto utile: permette infatti di analizzare efficacemente molte questioni complesse, considerando soltanto gli elementi essenziali e tralasciando gli aspetti secondari dal nostro punto di vista.

Oggetto della fisica matematica è lo studio delle relazioni fra questi elementi idealizzati, relazioni descritte mediante insiemi di funzioni di più variabili indipendenti.

§2. Le equazioni più semplici della fisica matematica

I legami e i rapporti elementari fra le grandezze fisiche sono espressi dalle leggi della meccanica e della fisica. Benché già molto varie, queste relazioni ne generano altre, più complesse e ancora più varie, che si deducono dalle prime mediante ragionamenti matematici. Le leggi della meccanica e della fisica possono essere scritte in linguaggio matematico sotto forma di equazioni alle derivate parziali, o anche di equazioni integrali, che collegano fra loro funzioni incognite. Per chiarire il significato di questa affermazione, consideriamo alcuni esempi di equazioni della fisica matematica.

Equazioni di conservazione della massa e dell'energia termica.

Esprimiamo in forma matematica le leggi fisiche fondamentali che governano i moti di un mezzo.

1. Esprimiamo anzitutto la legge di conservazione della materia contenuta in un volume qualsiasi Ω, che delimitiamo idealmente nello spazio e manteniamo fisso. A tale scopo dobbiamo calcolare la massa della materia contenuta in questo volume. La massa $M_\Omega(t)$ è espressa dall'integrale

$$M_\Omega(t) = \iiint_\Omega \rho(x, y, z, t)\, dx\, dy\, dz.$$

Questa massa, naturalmente, non sarà costante: in un processo oscillatorio la densità in ciascun punto varia, poiché le particelle di materia, oscillando, entrano nel volume oppure ne escono. La velocità di variazione della massa si trova derivando rispetto al tempo ed è data dall'integrale

$$\frac{dM_\Omega}{dt} = \iiint_\Omega \frac{\partial \rho}{\partial t}\, dx\, dy\, dz.$$

La velocità di variazione della massa contenuta nel volume può essere calcolata anche in un altro modo. Possiamo esprimere la quantità di materia che attraversa ogni secondo la superficie S delimitante il volume Ω, contando con segno negativo la materia che esce da Ω. Consideriamo a tale scopo un elemento ds della superficie S, sufficientemente piccolo da poter essere ritenuto piano e da poter attribuire lo stesso spostamento a tutti i suoi punti. Seguiamo lo spostamento dei punti di questo elemento nell'intervallo da t a $t + dt$. Calcoliamo anzitutto il vettore

$$\boldsymbol{v} = \frac{d\boldsymbol{u}}{dt},$$

che rappresenta la velocità di ciascuna particella. Nel tempo *dt* le particelle situate su *ds* si spostano del vettore *v dt* e raggiungono la posizione ds_1; la posizione *ds* viene invece occupata dalle particelle che prima si trovavano in ds_2 (figura 1). Durante questo intervallo attraversa quindi la superficie la colonnina di materia che inizialmente si trovava fra ds_2 e *ds*. La sua altezza è *v dt* cos (*n*, *v*), dove *n* denota la normale esterna alla superficie; il volume della colonnina è pertanto

$$v \cos (\boldsymbol{n}, \boldsymbol{v})\, ds\, dt,$$

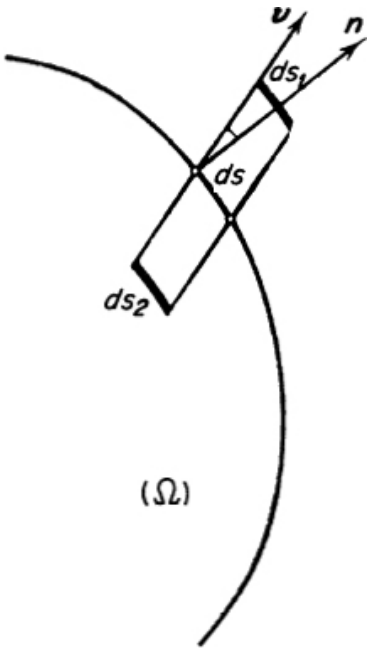


Figura 1.

e la sua massa è

$$\rho v \cos (\boldsymbol{n}, \boldsymbol{v})\, ds\, dt.$$

Sommando tutti questi piccoli contributi, otteniamo per la quantità di materia che esce dal volume nel tempo *dt* l'espressione

$$\iint_S \rho v \cos (\boldsymbol{n}, \boldsymbol{v})\, ds\, dt.$$

Nei punti in cui la velocità è diretta verso l'interno di *Ω*, il coseno è negativo: nell'integrale la materia entrante in *Ω* viene quindi contata con segno negativo. Il prodotto della velocità del mezzo per la sua densità si chiama densità di flusso. Il vettore densità di flusso della massa è $q = \rho v$.

Per trovare la portata di materia uscente dal volume *Ω* basta dividere questa espressione per *dt*; otteniamo così

$$\iint_S \rho v_n \, ds = \iint_S q_n \, ds,$$

dove

$$v_n = v \cos(n, v), \quad q_n = q \cos(n, q).$$

La componente normale del vettore v può essere sostituita con la sua espressione mediante le componenti dei vettori v e n lungo gli assi coordinati. Dalla geometria analitica sappiamo che

$$v_n = v \cos(n, v) = v_x \cos(n, x) + v_y \cos(n, y) + v_z \cos(n, z),$$

possiamo dunque riscrivere l'espressione della portata nella forma

$$\iint_S \rho(v_x \cos(n, x) + v_y \cos(n, y) + v_z \cos(n, z)) \, ds.$$

Per la legge di conservazione della materia, i due modi di calcolare la variazione della quantità di materia devono dare lo stesso risultato: ogni variazione della massa contenuta in Ω può avvenire soltanto per ingresso o uscita di materia attraverso la superficie S.

Uguagliando quindi la velocità di variazione della quantità di materia contenuta nel volume alla portata entrante, otteniamo

$$\begin{aligned} &\iiint_\Omega \frac{\partial \rho}{\partial t} \, dx \, dy \, dz \\ &\quad = -\iint_S [\rho v_x \cos(n, x) + \rho v_y \cos(n, y) + \rho v_z \cos(n, z)] \, ds \\ &\quad = -\iint_S [q_x \cos(n, x) + q_y \cos(n, y) + q_z \cos(n, z)] \, ds. \end{aligned}$$

Questa relazione integrale, come abbiamo detto, vale per ogni volume Ω. Si chiama «equazione di continuità».

L'integrale a secondo membro dell'ultima equazione può essere trasformato in un integrale di volume mediante la formula di Ostrogradskij. Questa formula, ricavata nel capitolo II, dà

$$\begin{aligned} &\iint_S (\rho v_x \cos(n, x) + \rho v_y \cos(n, y) + \rho v_z \cos(n, z)) \, ds \\ &\qquad = \iiint_\Omega \left[\frac{\partial(\rho v_x)}{\partial x} + \frac{\partial(\rho v_y)}{\partial y} + \frac{\partial(\rho v_z)}{\partial z}\right] d\Omega. \end{aligned}$$

Ne segue che

$$\iiint_\Omega \left[\frac{\partial \rho}{dt} + \frac{\partial(\rho v_x)}{\partial x} + \frac{\partial(\rho v_y)}{\partial y} + \frac{\partial(\rho v_z)}{\partial z}\right] d\Omega = 0.$$

Otteniamo così il seguente risultato: l'integrale della funzione

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v_x)}{\partial x} + \frac{\partial(\rho v_y)}{\partial y} + \frac{\partial(\rho v_z)}{\partial z} \quad \text{or} \quad \frac{\partial \rho}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z}$$

Nella formula: «or» = «oppure».

su ogni volume Ω è uguale a zero. Ma ciò è possibile soltanto se la funzione è identicamente nulla. Si ottiene dunque l'equazione di continuità in forma differenziale

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v_x)}{\partial x} + \frac{\partial(\rho v_y)}{\partial y} + \frac{\partial(\rho v_z)}{\partial z} = 0. \qquad (1)$$

L'equazione (1) è un esempio tipico della formulazione di una legge fisica nel linguaggio delle equazioni alle derivate parziali.

2. Consideriamo un altro problema dello stesso genere: la conduzione del calore.

In ogni mezzo le cui particelle sono animate da agitazione termica, il calore fluisce da alcuni punti ad altri. Questo flusso attraversa ogni elemento di superficie *ds* situato nel mezzo. Si può mostrare che il processo è descritto quantitativamente da un'unica grandezza vettoriale, il vettore densità di flusso termico, che indichiamo con τ. La quantità di calore che attraversa ogni secondo un elemento di area *ds* è allora $\tau_n\, ds$, così come $q_n\, ds$ esprimeva la quantità di materia che attraversava ogni secondo quell'area. Al posto del vettore densità di flusso di materia $q = \rho v$ abbiamo ora il vettore densità di flusso termico τ.

Nello stesso modo in cui abbiamo ottenuto l'equazione di continuità, che per il moto di un liquido esprime la conservazione della massa, possiamo ricavare una nuova equazione alle derivate parziali che esprime la conservazione dell'energia.

La densità volumica di energia termica Q in un dato punto può essere espressa dalla formula

$$Q = CT,$$

dove C è la capacità termica per unità di volume e T la temperatura.

È facile stabilire l'equazione

$$C\frac{\partial T}{\partial t} + \frac{\partial \tau_x}{\partial x} + \frac{\partial \tau_y}{\partial y} + \frac{\partial \tau_z}{\partial z} = 0. \qquad (2)$$

La derivazione è identica a quella dell'equazione di continuità, sostituendo alla «densità» la «densità di energia termica» e al flusso di massa il flusso di calore. Qui abbiamo supposto che nel mezzo non vi sia produzione di energia termica. Se però è presente una sorgente di calore, l'equazione (2) del bilancio termico deve essere modificata. Indicando con q la densità di produzione della sorgente, ossia l'energia termica prodotta per unità di volume in un secondo, l'equazione di conservazione dell'energia termica assume la forma più generale

$$C\frac{\partial T}{\partial t}+\frac{\partial \tau_x}{\partial x}+\frac{\partial \tau_y}{\partial y}+\frac{\partial \tau_z}{\partial z}=q. \qquad (3)$$

3. Un'altra equazione dello stesso tipo di quella di continuità si può ricavare derivando la (1) rispetto al tempo. Facciamolo per le piccole oscillazioni di un gas intorno a una posizione di equilibrio. Supponiamo che le variazioni della densità siano piccole e che le quantità $\partial\rho/\partial x$, $\partial\rho/\partial y$, $\partial\rho/\partial z$ e $\partial\rho/\partial t$ siano abbastanza piccole da poter trascurare i loro prodotti con v_x, v_y e v_z. Allora

$$\frac{\partial \rho}{\partial t}+\rho\left(\frac{\partial v_x}{\partial x}+\frac{\partial v_y}{\partial y}+\frac{\partial v_z}{\partial z}\right)=0.$$

Derivando rispetto al tempo e trascurando i prodotti di $\partial\rho/\partial t$ con $\partial v_x/\partial x$, $\partial v_y/\partial y$ e $\partial v_z/\partial z$, otteniamo

$$\frac{\partial^2 \rho}{\partial t^2}+\rho\left[\frac{\partial\left(\frac{dv_x}{dt}\right)}{\partial x}+\frac{\partial\left(\frac{dv_y}{dt}\right)}{\partial y}+\frac{\partial\left(\frac{dv_z}{dt}\right)}{\partial z}\right]=0. \qquad (4)$$

Equazione del moto.

1. Un esempio importante dell'espressione di una legge fisica mediante un'equazione differenziale è dato dalle equazioni di equilibrio o di moto di un mezzo. Il mezzo sia costituito da particelle materiali che si muovono con velocità diverse. Come nel primo esempio, delimitiamo idealmente nello spazio un volume Ω, racchiuso dalla superficie S e occupato dalle particelle del mezzo, e scriviamo la seconda legge di Newton per queste particelle. La legge afferma che, per ogni moto del mezzo, la velocità di variazione della quantità di moto totale delle particelle considerate è uguale alla somma di tutte le forze che agiscono su di esse. Come è noto dalla meccanica, la quantità di moto è rappresentata dal vettore

$$\boldsymbol{P}=\iiint_{\Omega}\rho\boldsymbol{v}\,d\Omega.$$

Le particelle che occupano un piccolo volume $d\Omega$ con densità ρ riempiranno, dopo un tempo Δt, un nuovo volume $d\Omega'$ con densità ρ', ma la loro massa resterà invariata:

$$\rho'\,d\Omega'=\rho\,d\Omega.$$

Se nello stesso intervallo la velocità v assume il nuovo valore v', variando quindi di $\Delta v = v' - v$, la corrispondente variazione della quantità di moto sarà

$$\rho'\boldsymbol{v}'\,d\Omega'-\rho\boldsymbol{v}\,d\Omega=\rho\boldsymbol{v}'\,d\Omega-\rho\boldsymbol{v}\,d\Omega=\rho\,\Delta\boldsymbol{v}\,d\Omega,$$

oppure, per unità di tempo,

$$\rho\frac{\Delta\boldsymbol{v}}{\Delta t}\,d\Omega\approx\rho\frac{d\boldsymbol{v}}{dt}\,d\Omega.$$

Sommando su tutte le particelle del volume Ω, troviamo che la velocità di variazione della quantità di moto è

$$\iiint_{\Omega} \rho \frac{d\boldsymbol{v}}{dt}\, d\Omega$$

ossia

$$\iiint_{\Omega} \rho \frac{dv_x}{dt}\, d\Omega,\ \iiint_{\Omega} \rho \frac{dv_y}{dt}\, d\Omega,\ \iiint_{\Omega} \rho \frac{dv_z}{dt}\, d\Omega.$$

Qui le derivate dv_x/dt, dv_y/dt e dv_z/dt denotano la velocità di variazione delle componenti di v seguendo una determinata particella, non osservando un punto fisso dello spazio. Questo è il significato della notazione d/dt anziché $\partial/\partial t$. Come è noto, $d/dt = \partial/\partial t + v_x(\partial/\partial x) + v_y(\partial/\partial y) + v_z(\partial/\partial z)$.

Le forze agenti sul volume possono essere di due tipi: forze di volume, che agiscono su ogni particella del corpo, e forze di superficie, o sforzi, agenti sulla superficie S che delimita il volume. Le prime sono forze a distanza, le seconde forze di contatto.

Per illustrare queste osservazioni, supponiamo che il mezzo considerato sia un fluido ideale. Le forze di superficie agenti su un elemento ds hanno allora intensità $p\ ds$, dove p è la pressione, e verso opposto a quello della normale esterna.

Indicando con n il versore normale alla superficie S, la forza agente sull'elemento ds è

$$-p\boldsymbol{n}\ ds.$$

Se F denota il vettore delle forze esterne per unità di volume, la nostra equazione assume la forma

$$\iiint_{\Omega} \rho \frac{d\boldsymbol{v}}{dt}\, d\Omega = \iiint_{\Omega} \boldsymbol{F}\, d\Omega - \iint_{S} p\boldsymbol{n}\, ds.$$

Questa è l'equazione del moto in forma integrale. Come quella di continuità, può essere trasformata in forma differenziale. Otteniamo il sistema

$$\rho \frac{dv_x}{dt} + \frac{\partial p}{\partial x} = F_x,\ \rho \frac{dv_y}{dt} + \frac{\partial p}{\partial y} = F_y,\ \rho \frac{\partial v_z}{dt} + \frac{\partial p}{\partial z} = F_z. \qquad (5)$$

Questo sistema è la forma differenziale della seconda legge di Newton.

2. Un altro esempio caratteristico dell'applicazione delle leggi della meccanica in forma differenziale è l'equazione della corda vibrante. Una corda è un corpo elastico lungo e molto sottile, flessibile proprio per la sua estrema sottigliezza e solitamente mantenuto ben teso. Se immaginiamo di dividerla in due parti in un punto x, su ciascuna parte agisce una forza pari alla tensione, diretta secondo la tangente alla curva della corda.

Esaminiamo un breve tratto della corda. Indichiamo con $u(x, t)$ lo spostamento di un suo punto dalla posizione di equilibrio. Supponiamo che l'oscillazione avvenga in un piano e consista di spostamenti perpendicolari all'asse Ox; rappresentiamo graficamente lo spostamento in un dato istante (figura 2). Studiamo il tratto compreso fra i punti x_1 e x_2. Alle sue estremità agiscono due forze di intensità pari alla tensione T, dirette secondo le rispettive tangenti a $u(x, t)$.

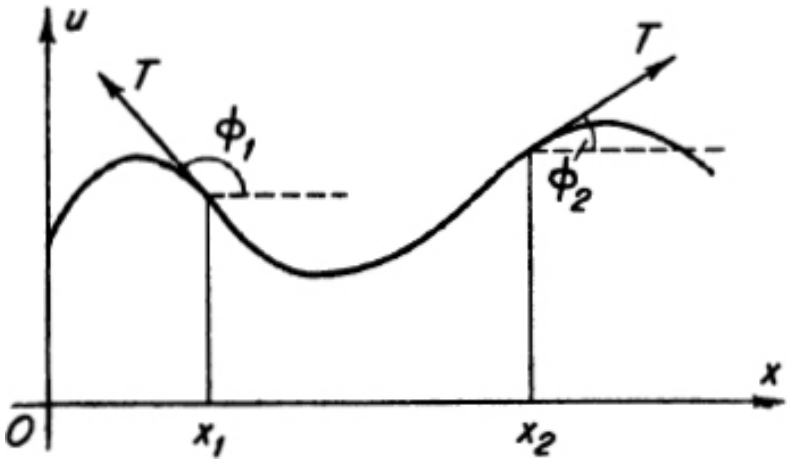


Figura 2.

Se il tratto è curvo, la risultante di queste due forze non è nulla. Secondo le leggi della meccanica, essa deve uguagliare la velocità di variazione della quantità di moto del tratto.

Sia ρ la massa per centimetro di lunghezza della corda. La velocità di variazione della quantità di moto è allora

$$\rho \int_{x_1}^{x_2} \frac{d^2u}{dt^2}\, dx.$$

Indicando con ϕ l'angolo fra la tangente alla corda e l'asse Ox, abbiamo

$$T \sin \phi_2 - T \sin \phi_1 = \int_{x_1}^{x_2} \rho \frac{\partial^2 u}{\partial t^2}\, dx.$$

È la consueta espressione della seconda legge della dinamica in forma integrale. È facile trasformarla in forma differenziale. Evidentemente

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x}(T \sin \phi).$$

Dai noti teoremi del calcolo differenziale si ricava facilmente il legame fra $T \sin \phi$ e la funzione incognita u:

$$\tan \phi = \frac{\partial u}{\partial x}, \sin \phi = \frac{\tan \phi}{\sqrt{1 + \tan^2 \phi}} = \frac{\partial u/\partial x}{\sqrt{1 + (\partial u/\partial x)^2}}$$

e, supponendo piccolo $\partial u/\partial x$, si ha

$$\sin\phi \approx \frac{\partial u}{\partial x}.$$

Quindi

$$T\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2}. \qquad (6)$$

Quest'ultima è l'equazione della corda vibrante in forma differenziale.

Forme fondamentali delle equazioni della fisica matematica.

Come già osservato, le varie equazioni alle derivate parziali che descrivono i fenomeni fisici formano di solito un sistema con più funzioni incognite. In moltissimi casi, però, è possibile sostituire il sistema con un'unica equazione, come mostrano esempi assai semplici.

Torniamo, per esempio, alle equazioni del moto considerate nel paragrafo precedente. Esse devono essere risolte insieme all'equazione di continuità. Esamineremo più avanti i metodi effettivi di soluzione.

1. Cominciamo dall'equazione del moto stazionario di un fluido ideale.

Tutti i moti possibili di un fluido si dividono in rotazionali e irrotazionali; questi ultimi si chiamano anche moti potenziali. Benché i moti irrotazionali siano casi particolari e, in generale, il moto di un liquido o di un gas sia sempre più o meno rotazionale, l'esperienza mostra che in molti casi esso è irrotazionale con grande approssimazione. Inoltre, nelle opportune ipotesi, si può dimostrare teoricamente che in un fluido a viscosità nulla un moto inizialmente irrotazionale rimane tale.

Per un moto potenziale esiste una funzione scalare $U(x, y, z, t)$, detta potenziale delle velocità, mediante la quale il vettore velocità v si esprime con le formule

$$v_x = \frac{\partial U}{\partial x}, \quad v_y = \frac{\partial U}{\partial y}, \quad v_z = \frac{\partial U}{\partial z}.$$

In tutti i casi finora studiati abbiamo incontrato sistemi di quattro equazioni con quattro funzioni incognite; in altri termini, un'equazione scalare e un'equazione vettoriale, contenenti una funzione scalare incognita e un campo vettoriale incognito. Di solito queste equazioni si possono riunire in un'unica equazione con una sola funzione incognita, ma tale equazione sarà del secondo ordine. Procediamo cominciando dal caso più semplice.

Per il moto potenziale di un fluido incomprimibile a densità costante, per il quale $\partial\rho/\partial t = 0$, abbiamo due gruppi di relazioni: l'equazione di continuità

$$\rho\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = 0$$

e le equazioni del moto potenziale

$$v_x = \frac{\partial U}{\partial x}, \quad v_y = \frac{\partial U}{\partial y}, \quad v_z = \frac{\partial U}{\partial z}.$$

Sostituendo nella prima equazione le espressioni della velocità date dalle seconde, otteniamo

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0. \qquad (7)$$

2. Anche il campo vettoriale del «flusso di calore» può essere espresso, mediante equazioni differenziali, in funzione di una sola grandezza scalare: la temperatura. È noto che il calore «fluisce» da un corpo caldo a uno freddo. Il vettore densità di flusso termico è dunque diretto in verso opposto al cosiddetto vettore gradiente della temperatura. È naturale supporre, e l'esperienza lo conferma, che in prima approssimazione il suo modulo sia direttamente proporzionale al modulo del gradiente della temperatura.

Le componenti del gradiente della temperatura sono

$$\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}.$$

Indicando con k il coefficiente di proporzionalità, otteniamo tre equazioni

$$\tau_x = -k \frac{\partial T}{\partial x}, \ \tau_y = -k \frac{\partial T}{\partial y}, \ \tau_z = -k \frac{\partial T}{\partial z}.$$

da risolvere insieme all'equazione di conservazione dell'energia termica

$$C \frac{\partial T}{\partial t} + \frac{\partial \tau_x}{\partial x} + \frac{\partial \tau_y}{\partial y} + \frac{\partial \tau_z}{\partial z} = q.$$

Sostituendo a τ_x, τ_y e τ_z le loro espressioni in funzione di T, otteniamo

$$C \frac{\partial T}{\partial t} = k \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \right) + q. \qquad (8)$$

3. Infine, per le piccole vibrazioni in un mezzo gassoso, come quelle sonore, l'equazione

$$\frac{\partial^2 \rho}{\partial t^2} + \rho \frac{\partial}{\partial x}\left(\frac{dv_x}{dt}\right) + \rho \frac{\partial}{\partial y}\left(\frac{dv_y}{dt}\right) + \rho \frac{\partial}{\partial z}\left(\frac{dv_z}{dt}\right) = 0$$

e le equazioni della dinamica (5) danno

$$\rho \frac{dv_x}{dt} + \frac{\partial p}{\partial x} = F_x, \ \rho \frac{dv_y}{dt} + \frac{\partial p}{dy} = F_y, \ \rho \frac{dv_z}{dt} + \frac{\partial p}{\partial z} = F_z,$$

e, supponendo assenti le forze esterne ($F_x = F_y = F_z = 0$), ricaviamo

$$\frac{\partial^2 p}{\partial t^2} = a^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right) \qquad (9)$$

Per ottenere questa equazione basta sostituire l'espressione delle accelerazioni nell'equazione di continuità ed eliminare la densità ρ mediante la legge di Boyle-Mariotte: $p = a^2\rho$.

Le equazioni (7), (8) e (9) sono tipiche di molti problemi della fisica matematica, oltre a quelli qui considerati. Lo studio approfondito di queste equazioni permette di comprendere numerose situazioni fisiche.

§3. Problemi ai valori iniziali e al contorno; unicità della soluzione

Per le equazioni alle derivate parziali, come per quelle ordinarie, accade, salvo rare eccezioni, che ogni equazione possieda infinite soluzioni particolari. Per risolvere un problema fisico concreto, cioè trovare una funzione incognita che soddisfi una certa equazione, occorre quindi saper scegliere la soluzione richiesta fra infinite soluzioni. A tale scopo, di solito, bisogna conoscere non soltanto l'equazione, ma anche alcune condizioni supplementari. Come abbiamo visto, le equazioni alle derivate parziali esprimono leggi elementari della meccanica o della fisica riferite a piccole particelle di un mezzo. Ma conoscere soltanto le leggi della meccanica non basta a prevedere lo svolgimento di un processo. Per esempio, per prevedere il moto dei corpi celesti, come si fa in astronomia, non basta conoscere la formulazione generale delle leggi di Newton: supponendo note le masse dei corpi, occorre conoscere anche lo stato iniziale del sistema, ossia le posizioni e le velocità dei corpi in un certo istante iniziale. Condizioni supplementari di questo genere si incontrano sempre nella soluzione dei problemi della fisica matematica.

I problemi della fisica matematica consistono dunque nel trovare soluzioni di equazioni alle derivate parziali che soddisfino determinate condizioni supplementari.

Le equazioni (7), (8) e (9) differiscono per struttura. Corrispondentemente diversi sono i problemi fisici che si possono risolvere mediante esse.

Le equazioni di Laplace e di Poisson; funzioni armoniche e unicità della soluzione dei problemi al contorno.

Analizziamo questi problemi più da vicino, cominciando dalle equazioni di Laplace e di Poisson. L'equazione di Poisson è[1]

$$\Delta u = -4\pi\rho,$$

dove ρ rappresenta di solito una densità. In particolare, ρ può annullarsi. Per $\rho \equiv 0$ si ottiene l'equazione di Laplace

$$\Delta u = 0.$$

È facile vedere che la differenza fra due soluzioni particolari qualsiasi u_1 e u_2 dell'equazione di Poisson è una funzione che soddisfa l'equazione di Laplace, ossia una funzione armonica. L'intera famiglia delle soluzioni dell'equazione di Poisson si riconduce così alla famiglia delle funzioni armoniche.

Se siamo riusciti a costruire anche una sola soluzione particolare u_0 dell'equazione di Poisson e definiamo una nuova funzione incognita w mediante

$$u = u_0 + w,$$

vediamo che w deve soddisfare l'equazione di Laplace; nello stesso modo determiniamo le corrispondenti condizioni al contorno per w. È dunque particolarmente importante studiare i problemi al contorno per l'equazione di Laplace.

Come accade molto spesso nei problemi matematici, la corretta formulazione di un problema per un'equazione della fisica matematica è suggerita direttamente dalla situazione pratica. Le condizioni supplementari che intervengono nella soluzione dell'equazione di Laplace derivano dalla formulazione fisica del problema.

Consideriamo, per esempio, l'instaurarsi di una temperatura stazionaria in un mezzo: la propagazione del calore in presenza di sorgenti costanti, situate all'interno o all'esterno del mezzo. In queste condizioni, con il trascorrere del tempo la temperatura raggiunta in ciascun punto diviene indipendente dal tempo. Per trovare la temperatura T in ogni punto dobbiamo quindi cercare quella soluzione dell'equazione

$$\frac{\partial T}{\partial t} = \Delta T + q,$$

dove q è la densità delle sorgenti termiche, che non dipende da t. Otteniamo

$$\Delta T + q = 0.$$

La temperatura nel mezzo soddisfa dunque l'equazione di Poisson. Se la densità delle sorgenti q è nulla, questa si riduce all'equazione di Laplace.

Per determinare la temperatura all'interno del mezzo, semplici considerazioni fisiche mostrano che bisogna conoscere anche ciò che avviene sul suo contorno.

È evidente che le leggi fisiche prima considerate per i punti interni di un corpo richiedono una formulazione diversa nei punti del contorno.

Nel problema della temperatura stazionaria possiamo assegnare la distribuzione della temperatura sul contorno, oppure il flusso di calore per unità di area della superficie, oppure, infine, una legge che colleghi temperatura e flusso termico.

Per la temperatura in un volume Ω, delimitato dalla superficie S, queste tre condizioni si scrivono:

$$T|_S = \phi(Q), \qquad (10)$$

oppure

$$\left.\frac{\partial T}{\partial n}\right|_S = \psi(Q), \qquad (10')$$

oppure, infine, nel caso più generale,

$$\alpha \left.\frac{\partial T}{\partial n}\right|_S + \beta T|_S = \chi(Q), \qquad (10'')$$

dove Q denota un punto qualsiasi della superficie S. Le condizioni della forma (10) si chiamano condizioni al contorno. Lo studio dell'equazione di Laplace o di Poisson con condizioni al contorno di uno di questi tipi mostra che, di regola e con le precisazioni indicate più avanti, la soluzione è determinata univocamente.

Per cercare una soluzione dell'equazione di Laplace o di Poisson, è quindi generalmente necessario e sufficiente assegnare una funzione arbitraria sul contorno del dominio.[2] Esaminiamo più da vicino l'equazione di Laplace. Mostreremo che una funzione armonica u, cioè una funzione che soddisfa tale equazione, è completamente determinata dai suoi valori sul contorno del dominio.

Stabiliamo anzitutto che una funzione armonica non può assumere all'interno del dominio valori maggiori del massimo valore sul contorno. Più precisamente, mostriamo che tanto il massimo assoluto quanto il minimo assoluto di una funzione armonica sono raggiunti sul contorno del dominio.

Ne segue subito che, se una funzione armonica ha valore costante sul contorno di un dominio Ω, essa è uguale alla stessa costante anche all'interno. Infatti, se il massimo e il minimo di una funzione coincidono con una medesima costante, la funzione è ovunque uguale a essa.

Dimostriamo ora che una funzione armonica non può avere all'interno valori superiori al massimo sul contorno o inferiori al minimo sul contorno. Osserviamo anzitutto che, se il laplaciano Δu di $u(x, y, z)$ è positivo in tutto il dominio, la funzione non può avere un massimo interno; se è negativo, non può avere un minimo interno. Infatti, in un punto di massimo u deve avere un massimo anche come funzione di ciascuna variabile separatamente, mantenendo fisse le altre. Ogni derivata seconda rispetto a una singola variabile deve quindi essere non positiva. Anche la loro somma sarebbe non positiva, mentre il laplaciano è positivo: una contraddizione. Analogamente, se la funzione ha un minimo in un punto interno, il suo laplaciano non può esservi negativo. Un laplaciano ovunque negativo esclude dunque un minimo interno.

Una funzione armonica può sempre essere modificata di una quantità arbitrariamente piccola in modo da ottenere un laplaciano positivo o negativo; basta aggiungerle la quantità

$$\pm \eta r^2 = \pm \eta(x^2 + y^2 + z^2),$$

dove η è una costante positiva arbitrariamente piccola.

Se il valore del massimo interno fosse strettamente maggiore di tutti i valori sul contorno, l'aggiunta di una quantità abbastanza piccola non eliminerebbe questa proprietà. Aggiungendo $+\eta r^2$ alla funzione armonica otterremmo allora una funzione con laplaciano positivo che avrebbe ancora un massimo interno, in contrasto con quanto appena dimostrato. Nessun valore interno può quindi superare il massimo sul contorno. Con un ragionamento analogo si dimostra che nessun valore interno può essere inferiore al minimo sul contorno.

Questo teorema ha un'importante conseguenza: due funzioni armoniche che coincidono sul contorno di un dominio coincidono anche in tutto il suo interno. La loro differenza, anch'essa armonica, si annulla infatti sul contorno e deve pertanto essere identicamente nulla all'interno.

I valori sul contorno determinano dunque completamente una funzione armonica. Si può dimostrare, benché qui non possiamo entrare nei dettagli, che, sotto opportune ipotesi di regolarità sul dominio e sui dati, esiste una funzione armonica che assume i valori assegnati arbitrariamente sul contorno.

È un po' più complesso studiare come la temperatura stazionaria di un corpo sia determinata dalla conoscenza del flusso termico attraverso ciascun elemento della superficie o di una legge che colleghi flusso e temperatura. Torneremo su alcuni aspetti della questione discutendo i metodi di soluzione dei problemi della fisica matematica.

Il problema al contorno per l'equazione del calore.

Una situazione del tutto diversa si presenta nel caso non stazionario. È fisicamente evidente che i valori della temperatura sul contorno o del flusso termico attraverso di esso non bastano da soli a determinare un'unica soluzione. Se però conosciamo anche la distribuzione della temperatura in un istante iniziale, il problema è determinato univocamente. Per determinare la soluzione dell'equazione di conduzione del calore (8), è dunque di solito necessario e sufficiente assegnare una funzione arbitraria $T_0(x, y, z)$, che descriva la distribuzione iniziale della temperatura, e una funzione arbitraria sul contorno del dominio. Come prima, quest'ultima può rappresentare la temperatura della superficie, il flusso termico attraverso ciascun suo elemento oppure una legge che colleghi flusso e temperatura.

Il problema si formula quindi così: cercare una soluzione dell'equazione (8) che soddisfi la condizione

$$T|_{t=0} = T_0(x, y, z) \qquad (11)$$

e una delle tre condizioni seguenti

$$T|_S = \phi(Q), \qquad (12)$$

$$\left.\frac{\partial T}{\partial n}\right|_S = \psi(Q), \qquad (12')$$

$$\alpha \frac{\partial T}{\partial n}\Big|_S + \beta T|_S = \chi(Q), \qquad (12'')$$

dove Q è un punto qualsiasi della superficie S.

La (11) si chiama condizione iniziale, mentre le (12) sono condizioni al contorno.

Non dimostreremo in dettaglio che ciascuno di questi problemi possiede una soluzione unica, ma stabiliremo l'unicità soltanto per il primo; inoltre considereremo il caso senza sorgenti termiche interne. Mostriamo che l'equazione

$$\Delta T = \frac{1}{a^2}\frac{\partial T}{\partial t}$$

con le condizioni

$$T|_{t=0} = T_0(x, y, z),$$
$$T|_S = \phi(Q)$$

può avere al più una soluzione.

La dimostrazione è molto simile a quella precedente per l'unicità della soluzione dell'equazione di Laplace. Mostriamo anzitutto che, se

$$\Delta T - \frac{1}{a^2}\frac{\partial T}{\partial t} < 0,$$

allora la funzione T, considerata nelle quattro variabili x, y, z e $t(0 \leqslant t \leqslant t_0)$, assume il minimo sul contorno del dominio Ω oppure al suo interno, ma in quest'ultimo caso necessariamente all'istante iniziale $t = 0$.

Altrimenti il minimo sarebbe raggiunto in un punto spaziale interno a un tempo positivo. Se tale tempo è interno all'intervallo considerato, tutte le derivate prime, compresa $\partial T/\partial t$, sono nulle; se il minimo si verifica al tempo finale $t = t_0$, la derivata temporale è non positiva. Inoltre tutte le derivate seconde rispetto a x, y e z sono non negative. Di conseguenza $\Delta T - (1/a^2)(\partial T/\partial t)$ è non negativo, il che nel nostro caso è impossibile.

Nello stesso modo si dimostra che, se $\Delta T - (1/a^2)(\partial T/\partial t) > 0$, la funzione T non può avere un massimo all'interno di Ω per $0 < t \leqslant t_0$.

Infine, se $\Delta T - (1/a^2)(\partial T/\partial t) = 0$, nessun valore interno per $0 < t \leqslant t_k$ può essere inferiore a tutti i valori iniziali e al contorno, né superiore a essi. Se, per esempio, vi fosse un minimo strettamente inferiore a tali valori, aggiungendo il termine $\eta(t - t_0)$ e considerando $T_1 = T + \eta(t - t_0)$, con η positivo abbastanza piccolo, conserveremmo un minimo interno a tempo positivo. Ma allora $\Delta T_1 - (1/a^2)(\partial T_1/\partial t)$ sarebbe negativo, il che è impossibile.

Analogamente si esclude un massimo strettamente superiore ai valori iniziali e al contorno nel dominio considerato.

Il massimo e il minimo assoluti della temperatura sono dunque raggiunti all'istante iniziale $t = 0$ oppure sul contorno S del mezzo. Se $T = 0$ sia inizialmente sia sul contorno, allora $T = 0$ identicamente in tutto l'interno per ogni $t \leqslant t_0$. Se due distribuzioni di temperatura T_1 e T_2 hanno gli stessi valori per $t = 0$ e sul contorno, la differenza $T_1 - T_2 = T$ soddisfa l'equazione del calore e si annulla inizialmente e sul contorno. Questa differenza è quindi ovunque nulla: le due distribuzioni coincidono.

Studiando più avanti i metodi di soluzione delle equazioni della fisica matematica, vedremo che i valori di T per $t = 0$ e il secondo membro di una delle (12) possono essere assegnati arbitrariamente, con le consuete condizioni di regolarità e compatibilità: la soluzione del problema esiste.

L'energia delle oscillazioni e il problema al contorno per l'equazione delle onde.

Consideriamo ora le condizioni che assicurano l'unicità della soluzione della terza equazione differenziale fondamentale, la (9).

Per semplicità consideriamo l'equazione della corda vibrante $\partial^2 u/\partial x^2 = (1/a^2)(\partial^2 u/\partial t^2)$, molto simile alla (9), dalla quale differisce soltanto per il numero delle variabili spaziali. A secondo membro compare $\partial^2 u/\partial t^2$, che esprime l'accelerazione di un punto arbitrario della corda. Il moto di un sistema meccanico nel quale le forze, e quindi le accelerazioni, sono espresse dalle coordinate dei corpi in movimento è completamente determinato dalle posizioni e dalle velocità iniziali di tutti i punti del sistema. Per l'equazione della corda vibrante è quindi naturale assegnare posizioni e velocità di tutti i punti all'istante iniziale:

$$u\,|_{t=0} = u_0(x)$$
$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = u_1(x).$$

Ma, come osservato, alle estremità della corda non si applicano le formule che esprimono le leggi della meccanica per i punti interni. A entrambe le estremità dobbiamo dunque assegnare condizioni supplementari. Per esempio, se la corda è fissata agli estremi nella posizione di equilibrio, avremo

$$u\,|_{x=0} = u\,|_{x=l} = 0.$$

Queste condizioni possono talvolta essere sostituite con altre più generali, ma ciò non modifica la sostanza del problema.

Il problema di trovare le soluzioni richieste dell'equazione (9) è analogo. Per determinarle si assegnano abitualmente le condizioni

$$p\,|_{t=0} = \phi_0(x, y, z),$$
$$\left.\frac{\partial p}{\partial t}\right|_{t=0} = \phi_1(x, y, z), \qquad (13)$$

e una delle «condizioni al contorno»

$$p\,|_S = \phi(Q), \qquad (14)$$

$$\frac{\partial p}{\partial n}\Big|_S = \psi(Q), \qquad (14')$$

$$\alpha \frac{\partial p}{\partial n}\Big|_S + \beta p|_S = \chi(Q).^* \qquad (14'')$$

La differenza rispetto al caso precedente sta semplicemente nel fatto che, al posto dell'unica condizione iniziale (11), abbiamo le due condizioni (13).[3]

Le (14) esprimono evidentemente le leggi fisiche per le particelle situate sul contorno del volume considerato.

Ometteremo la dimostrazione generale che le condizioni (13), insieme a una delle (14), determinano univocamente la soluzione. Mostreremo soltanto l'unicità per una delle condizioni (14).

Supponiamo che una funzione u soddisfi l'equazione

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{a^2}\frac{\partial^2 u}{\partial t^2},$$

con le condizioni iniziali

$$u\,|_{t=0} = 0, \frac{\partial u}{\partial t}\Big|_{t=0} = 0$$

e la condizione al contorno

$$\frac{\partial u}{\partial n}\Big|_S = 0.$$

Sarebbe altrettanto facile trattare il caso $u|_S = 0$.

Mostreremo che, sotto queste condizioni, la funzione u deve essere identicamente nulla.

Per dimostrarlo non bastano gli argomenti usati prima per l'unicità nei primi due problemi. Qui possiamo però ricorrere all'interpretazione fisica.

Ci servirà una sola legge fisica: la «legge di conservazione dell'energia». Per semplicità torniamo alla corda vibrante, lo spostamento dei cui punti $u(x, t)$ soddisfa l'equazione

$$T\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2}.$$

L'energia cinetica dell'elemento di corda compreso fra x e $x + dx$ è espressa da

$$\frac{1}{2}\left(\frac{\partial u}{\partial t}\right)^2 \rho\, dx.$$

Oltre all'energia cinetica, la corda spostata dalla posizione di equilibrio possiede energia potenziale, dovuta all'aumento della sua lunghezza rispetto alla posizione rettilinea. Calcoliamola considerando l'elemento compreso fra x e $x + dx$. Esso è inclinato rispetto all'asse Ox e la sua lunghezza è approssimativamente

$$\sqrt{(dx)^2 + \left(\frac{\partial u}{\partial x}\,dx\right)^2};$$

perciò il suo allungamento è

$$\sqrt{1 + \left(\frac{\partial u}{\partial x}\right)^2}\,dx - dx \approx \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2 dx.$$

Moltiplicando l'allungamento per la tensione T, troviamo l'energia potenziale dell'elemento allungato:

$$\frac{1}{2}T\left(\frac{\partial u}{\partial x}\right)^2 dx.$$

L'energia totale della corda di lunghezza l si ottiene sommando le energie cinetica e potenziale su tutti i suoi punti:

$$E = \frac{1}{2}\int_0^l \left[T\left(\frac{\partial u}{\partial x}\right)^2 + \rho\left(\frac{\partial u}{\partial t}\right)^2\right] dx.$$

Se le forze agenti alle estremità non compiono lavoro, in particolare se le estremità sono fisse, l'energia totale deve rimanere costante:

$$E = \text{const.}$$

Questa espressione della legge di conservazione dell'energia è una conseguenza matematica delle equazioni fondamentali della meccanica e può essere dedotta da esse. Avendo già scritto le leggi del moto mediante l'equazione differenziale della corda vibrante e le condizioni agli estremi, possiamo darne la seguente dimostrazione matematica. Derivando E rispetto al tempo, le regole generali danno

$$\frac{dE}{dt} = \int_0^l \left(T\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x\,\partial t} + \rho\frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial t^2}\right) dx.$$

Usando l'equazione delle onde (6) e sostituendo $\rho(\partial^2 u/\partial t^2)$ con $T(\partial^2 u/\partial x^2)$, otteniamo dE/dt nella forma

$$\frac{dE}{dt} = \int_0^l T\left[\left(\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x\,\partial t}\right) + \frac{\partial u}{\partial t}\frac{\partial^2 u}{\partial x^2}\right] dx$$
$$= \int_0^l T\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\frac{\partial u}{\partial t}\right) dx = T\frac{\partial u}{\partial x}\frac{\partial u}{\partial t}\Big|_{x=l} - T\frac{\partial u}{\partial x}\frac{\partial u}{\partial t}\Big|_{x=0}.$$

Se a ciascuno dei due estremi si annulla $\partial u/\partial x$ oppure lo spostamento u è mantenuto nullo, allora

$$\frac{dE}{dt} = 0,$$

e quindi E è costante.

L'equazione delle onde (9) può essere trattata nello stesso modo, dimostrando che anche per essa vale la conservazione dell'energia. Se p soddisfa la (9) e la condizione

$$p\,|_S = 0 \quad \text{or} \quad \frac{\partial p}{\partial n}\Big|_S = 0,$$

Nella formula: «or» = «oppure».

allora la quantità

$$E = \iiint \left[\left(\frac{\partial p}{\partial x}\right)^2 + \left(\frac{\partial p}{\partial y}\right)^2 + \left(\frac{\partial p}{\partial z}\right)^2 + \frac{1}{a^2}\left(\frac{\partial p}{\partial t}\right)^2\right] dx\,dy\,dz$$

non dipende da t.

Se all'istante iniziale l'energia totale delle oscillazioni è nulla, rimarrà sempre nulla, e ciò è possibile soltanto in assenza di moto. Se il problema di integrazione dell'equazione delle onde con condizioni iniziali e al contorno avesse due soluzioni p_1 e p_2, la differenza $v = p_1 - p_2$ soddisferebbe la stessa equazione con condizioni aventi secondo membro nullo, cioè condizioni omogenee.

Calcolando l'«energia» dell'oscillazione descritta da v, troveremmo $E(v) = 0$ all'istante iniziale. Essa sarebbe quindi sempre nulla e, tenendo conto dello spostamento iniziale nullo, anche v sarebbe identicamente nulla. Le due soluzioni p_1 e p_2 coinciderebbero: la soluzione è unica.

Abbiamo così verificato l'unicità per le tre formulazioni considerate; l'esistenza verrà affrontata attraverso i metodi di soluzione.

Abbiamo inoltre scoperto alcune proprietà molto semplici delle soluzioni. Per esempio, le soluzioni dell'equazione di Laplace possiedono la seguente proprietà di massimo: i loro valori massimo e minimo sono raggiunti sul contorno del dominio di definizione.

Le funzioni che descrivono la distribuzione del calore possiedono una proprietà di massimo di tipo diverso. Ogni massimo o minimo locale della temperatura tende progressivamente a

smussarsi col tempo. La temperatura in un punto può aumentare o diminuire soltanto se è rispettivamente più bassa o più alta di quella dei punti vicini. Col trascorrere del tempo la temperatura si uniforma: le disuguaglianze si attenuano grazie al passaggio del calore dalle zone calde a quelle fredde.

Nulla di simile avviene nella propagazione delle oscillazioni qui considerate. Esse non si attenuano né si uniformano, poiché la somma dell'energia cinetica e di quella potenziale deve rimanere costante in ogni istante.

§4. La propagazione delle onde

Le proprietà delle oscillazioni si possono illustrare con grande chiarezza mediante esempi semplicissimi. Consideriamo due casi caratteristici.

Il primo esempio è l'equazione della corda vibrante

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{a^2}\frac{\partial^2 u}{\partial t^2}. \qquad (15)$$

Come si può verificare, questa equazione ammette due famiglie di soluzioni particolari della forma

$$u_1 = \phi_1(x - at), \quad u_2 = \phi_2(x + at),$$

dove ϕ_1 e ϕ_2 sono funzioni arbitrarie derivabili due volte.

Derivando direttamente è facile mostrare che u_1 e u_2 soddisfano la (15). Si può dimostrare che

$$u = u_1 + u_2$$

è la soluzione generale dell'equazione.

La forma delle oscillazioni descritte da u_1 e u_2 è di notevole interesse. Per esaminarla nel modo più comodo, compiamo idealmente il seguente esperimento. L'osservatore della corda vibrante non resti fermo, ma si muova lungo l'asse *Ox* con velocità *a*. Per lui la posizione di un punto della corda sarà definita da un sistema di coordinate mobile. Indichiamo con ξ la coordinata corrispondente a *x* in questo sistema. Il valore $\xi = 0$ corrisponderà evidentemente, in ogni istante, a $x = at$. È quindi chiaro che

$$\xi = x - at.$$

Una funzione arbitraria $u(x, t)$ può essere rappresentata nella forma

$$u(x, t) = \phi(\xi, t).$$

Per la soluzione u_1 avremo

$$u_1(x, t) = \phi_1(\xi),$$

In questo sistema di coordinate la soluzione $u_1(x, t)$ risulta dunque indipendente dal tempo. Per l'osservatore che si muove con velocità a la corda appare come una curva immobile; per l'osservatore fermo, invece, lungo la corda si propaga un'onda nella direzione dell'asse Ox con velocità a.

Analogamente, la soluzione $u_2(x, t)$ rappresenta un'onda che viaggia nel verso opposto con velocità di modulo a. Se la corda è infinita, entrambe le onde si propagano indefinitamente. Muovendosi in direzioni opposte, la loro sovrapposizione può dare alla corda forme assai singolari: lo spostamento risultante può aumentare in certi istanti e diminuire in altri.

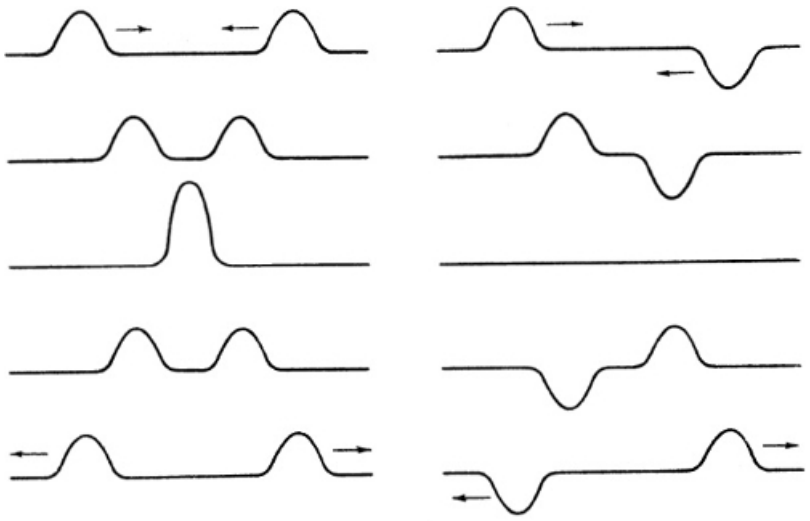


Figura 3.

Se u_1 e u_2, giungendo nello stesso punto da lati opposti, hanno lo stesso segno, si rafforzano; se hanno segno opposto, si contrastano. La figura 3 mostra alcune posizioni successive della corda per due particolari perturbazioni. Inizialmente le onde si avvicinano indipendentemente, poi cominciano a sovrapporsi. Nel secondo caso della figura si ha un istante di completo annullamento dello spostamento, dopo il quale le onde tornano a separarsi.

Un altro esempio che si presta facilmente a uno studio qualitativo è la propagazione delle onde nello spazio.

L'equazione

$$\Delta u = \frac{1}{a^2}\frac{\partial^2 u}{\partial t^2} \qquad (16)$$

ricavata in precedenza ammette due famiglie di soluzioni particolari della forma

$$u_1 = \frac{1}{r}\phi_1(r - at), \qquad u_2 = \frac{1}{r}\phi_2(r + at), \qquad (17)$$

dove r è la distanza del punto considerato dall'origine, $r^2 = x^2 + y^2 + z^2$, e ϕ_1, ϕ_2 sono funzioni arbitrarie derivabili due volte.

La verifica che u_1 e u_2 sono soluzioni richiederebbe un certo spazio e viene qui omessa.

La forma delle onde descritte da queste soluzioni è, in generale, simile a quella delle onde sulla corda. Trascurando per un momento il fattore $1/r$ a secondo membro, la prima soluzione rappresenta un'onda che viaggia verso valori crescenti di r. Essa è sfericamente simmetrica: ha lo stesso valore in tutti i punti alla stessa distanza r dall'origine.

Il fattore $1/r$ fa sì che l'ampiezza dell'onda sia inversamente proporzionale alla distanza dall'origine. Una simile oscillazione si chiama onda sferica divergente. Un'immagine intuitiva è offerta dai cerchi che si allargano sulla superficie dell'acqua quando vi si getta un sasso, con la differenza che in quel caso le onde sono circolari anziché sferiche.

Anche la seconda soluzione della (17) è molto interessante: è un'onda convergente, che viaggia verso l'origine. Nella singola espressione radiale il fattore $1/r$ può far crescere senza limite l'ampiezza all'avvicinarsi all'origine. La concentrazione della perturbazione in un punto può quindi produrre, pur partendo da oscillazioni piccole, un effetto assai intenso.

§5. Metodi di costruzione delle soluzioni

Sulla possibilità di scomporre una soluzione in soluzioni più semplici.

Le soluzioni dei problemi della fisica matematica formulati in precedenza si possono ricavare con procedimenti diversi, adatti ai singoli problemi. Alla base di questi metodi vi è però un'idea comune. Come abbiamo visto, le equazioni considerate per piccole perturbazioni sono lineari rispetto alle funzioni incognite e alle loro derivate. Anche le condizioni iniziali e al contorno sono lineari.

La differenza fra due soluzioni della stessa equazione soddisfa l'equazione ottenuta annullando il secondo membro. Questa si chiama equazione omogenea associata. Per esempio, all'equazione di Poisson $\Delta u = -4\pi\rho$ è associata l'equazione omogenea di Laplace $\Delta u = 0$.

Se le due soluzioni soddisfano anche le stesse condizioni al contorno, la loro differenza soddisfa la corrispondente condizione omogenea: sul contorno l'espressione considerata è nulla.

L'intera famiglia delle soluzioni dell'equazione con le condizioni al contorno assegnate si ottiene dunque sommando a una soluzione particolare che soddisfi le condizioni non omogenee tutte le possibili soluzioni dell'equazione omogenea con condizioni al contorno omogenee, ma non necessariamente con le condizioni iniziali assegnate.

Le soluzioni di un'equazione omogenea con condizioni al contorno omogenee possono essere sommate o moltiplicate per costanti senza cessare di essere soluzioni.

Se una soluzione dell'equazione omogenea con condizioni omogenee dipende da un parametro, anche l'integrazione rispetto al parametro, quando lecita, dà una soluzione dello stesso tipo. Questi fatti sono alla base del più importante metodo per risolvere i problemi lineari della fisica matematica: il metodo di sovrapposizione.

Si cerca la soluzione nella forma

$$u = u_0 + \sum_k u_k ,$$

dove u_0 è una soluzione particolare che soddisfa le condizioni al contorno, ma non necessariamente quelle iniziali, mentre le u_k sono soluzioni dell'equazione omogenea associata con le corrispondenti condizioni al contorno omogenee. Se l'equazione e le condizioni al contorno sono già omogenee, si può cercare la soluzione nella forma

$$u = \sum u_k .$$

Per poter soddisfare condizioni iniziali arbitrarie mediante la scelta delle soluzioni particolari u_k, dobbiamo disporre di una famiglia sufficientemente ricca di queste soluzioni.

Il metodo di separazione delle variabili.

Per costruire la famiglia di soluzioni necessaria esiste un procedimento chiamato metodo di separazione delle variabili, o metodo di Fourier.

Esaminiamolo, per esempio, nella soluzione del problema

$$\Delta u = \frac{\partial^2 u}{\partial t^2} , \qquad (18)$$

$$u|_S = 0, \quad u|_{t=0} = f_0(x, y, z), \quad u_t|_{t=0} = f_1(x, y, z).$$

Cercando una soluzione particolare, supponiamo anzitutto che la funzione u soddisfi la condizione al contorno $u|_S = 0$ e si possa esprimere come prodotto di due funzioni: una dipendente soltanto dal tempo t, l'altra soltanto dalle variabili spaziali:

$$u(x, y, z, t) = U(x, y, z)\, T(t).$$

Sostituendo questa forma nell'equazione, otteniamo

$$T(t)\, \Delta U = T''(t)\, U.$$

Dividendo entrambi i membri per TU, si ha

$$\frac{T''}{T} = \frac{\Delta U}{U} .$$

Il secondo membro dipende soltanto dalle variabili spaziali, mentre il primo è indipendente da esse. L'uguaglianza può quindi valere soltanto se entrambi i membri sono uguali alla stessa costante. Si ottiene il sistema di due equazioni

$$\frac{T''}{T} = -\lambda_k^2, \quad \frac{\Delta U}{U} = -\lambda_k^2.$$

La costante a secondo membro è indicata con $-\lambda_k^2$ per sottolineare che è negativa, come si può dimostrare rigorosamente. L'indice k ricorda che esistono infiniti valori possibili di $-\lambda_k^2$, ai quali corrispondono soluzioni che formano un sistema di funzioni completo in un preciso senso.

Eliminando i denominatori nelle due equazioni, otteniamo

$$T'' + \lambda_k^2 T = 0; \quad \Delta U + \lambda_k^2 U = 0.$$

La prima ha, come sappiamo, la semplice soluzione

$$T = A_k \cos \lambda_k t + B_k \sin \lambda_k t,$$

dove A_k e B_k sono costanti arbitrarie. L'espressione si può semplificare introducendo un angolo ausiliario ϕ. Abbiamo

$$\frac{A_k}{\sqrt{A_k^2 + B_k^2}} = \sin \phi_k, \quad \frac{B_k}{\sqrt{A_k^2 + B_k^2}} = \cos \phi_k, \quad \sqrt{A_k^2 + B_k^2} = M_k.$$

e quindi

$$T = \sqrt{A_k^2 + B_k^2} \sin (\lambda_k t + \phi_k) = M_k \sin (\lambda_k t + \phi_k).$$

La funzione T rappresenta un'oscillazione armonica di pulsazione λ_k, con uno sfasamento pari all'angolo ϕ_k.

Più difficile e interessante è trovare una soluzione dell'equazione

$$\Delta U + \lambda_k^2 U = 0 \qquad (19)$$

con condizioni al contorno omogenee assegnate, per esempio

$$U|_S = 0$$

dove S è il contorno del volume Ω considerato, oppure con un'altra condizione omogenea. Per i valori ammissibili del parametro, la soluzione non è sempre facile da esprimere come combinazione finita di funzioni note, ma può essere determinata con la precisione desiderata.

L'equazione $\Delta U + \lambda_k^2 U = 0$ con la condizione $U|_S = 0$ ha anzitutto la soluzione evidente $U \equiv 0$, che è banale e del tutto inutile per i nostri scopi. Se i numeri λ_k sono scelti a caso, in

generale il problema non ha altre soluzioni. Esistono però valori di λ_k per i quali l'equazione ammette una soluzione non banale.

I possibili valori della costante λ_k^2 sono determinati dalla richiesta che la (19) ammetta una soluzione non banale, cioè non identicamente nulla, che soddisfi $U|_S = 0$. Da questa richiesta segue anche che i numeri indicati con $-\lambda_k^2$ sono negativi.

Per ogni valore ammissibile di λ_k nella (19) troviamo almeno una funzione U_k. Possiamo così costruire una soluzione particolare dell'equazione delle onde (18) nella forma

$$u_k = M_k \sin(\lambda_k t + \phi_k)\, U_k(x, y, z).$$

Una tale soluzione si chiama oscillazione propria del volume considerato. La costante λ_k è la pulsazione propria e la funzione $U_k(x, y, z)$ ne determina la forma. Questa funzione si chiama abitualmente autofunzione, o funzione propria. In ogni istante, u_k, considerata come funzione delle variabili spaziali, differisce da U_k soltanto per un fattore di scala.

Non abbiamo spazio per dimostrare in dettaglio le numerose e notevoli proprietà delle oscillazioni proprie e delle autofunzioni; ci limiteremo quindi a elencarne alcune.

La prima proprietà consiste nel fatto che, per ogni volume limitato considerato nelle consuete ipotesi di regolarità, esiste una successione numerabile di pulsazioni proprie. Queste tendono all'infinito al crescere di k.

Un'altra proprietà si chiama ortogonalità: l'integrale sul dominio Ω del prodotto di autofunzioni corrispondenti a valori diversi di λ_k è uguale a zero.[4]

$$\iiint_\Omega U_k(x, y, z)\, U_j(x, y, z)\, dx\, dy\, dz = 0 \quad (j \neq k).$$

Per $j = k$ supporremo

$$\iiint_\Omega U_k(x, y, z)^2\, dx\, dy\, dz = 1.$$

È sempre possibile ottenere questa normalizzazione moltiplicando $U_k(x, y, z)$ per una costante opportuna: ciò non cambia il fatto che la funzione soddisfi la (19) e la condizione $U|_S = 0$.

Infine, la terza proprietà consiste nel fatto che, includendo tutte le autofunzioni necessarie e senza omettere alcun valore di λ_k, possiamo rappresentare con precisione arbitraria una funzione $f(x, y, z)$ del tutto arbitraria, purché soddisfi $f|_S = 0$ e abbia derivate prime e seconde continue. Una tale funzione può essere rappresentata dalla serie convergente

$$f(x, y, z) = \sum_{k=1}^{\infty} C_k U_k(x, y, z). \qquad (20)$$

La terza proprietà consente dunque, in linea di principio, di sviluppare f in serie di autofunzioni del problema; la seconda permette di calcolare tutti i coefficienti. Moltiplicando infatti entrambi i membri della (20) per $U_j(x, y, z)$ e integrando su Ω, otteniamo

$$\iiint_{\Omega} f(x, y, z)\, U_j(x, y, z)\, dx\, dy\, dz = \sum_{k=1}^{\infty} C_k \iiint U_k(x, y, z)\, U_j(x, y, z)\, dx\, dy\, dz.$$

Nella somma a secondo membro tutti i termini con $k \neq j$ si annullano per l'ortogonalità, mentre il fattore che moltiplica C_j è uguale a uno. Di conseguenza

$$C_j = \iiint_{\Omega} f(x, y, z)\, U_j(x, y, z)\, dx\, dy\, dz.$$

Queste proprietà delle oscillazioni proprie permettono ora di risolvere il problema generale delle oscillazioni con condizioni iniziali arbitrarie.

A tale scopo cerchiamo una soluzione nella forma

$$u = \sum U_k(x, y, z)\,(A_k \cos \lambda_k t + B_k \sin \lambda_k t) \qquad (21)$$

e scegliamo le costanti A_k e B_k in modo che

$$u\,|_{t=0} = f_0(x, y, z),$$
$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = f_1(x, y, z).$$

Ponendo $t = 0$ a secondo membro della (21), i termini con il seno si annullano e $\cos \lambda_k t$ diventa uno. Avremo quindi

$$f_0(x, y, z) = \sum_{k=1}^{\infty} A_k U_k(x, y, z).$$

La terza proprietà assicura la possibilità di una simile rappresentazione; dalla seconda otteniamo

$$A_k = \iiint_{\Omega} f_0(x, y, z)\, U_k(x, y, z)\, dx\, dy\, dz.$$

Analogamente, derivando la (21) rispetto a t e ponendo $t = 0$, abbiamo

$$\left.\frac{\partial u}{\partial t}\right|_{t=0} = f_1(x, y, z) = \sum_{k=1}^{\infty} \lambda_k (B_k \cos \lambda_k t - A_k \sin \lambda_k t)\,|_{t=0} U_k(x, y, z)$$

$$= \sum_{k=1}^{\infty} \lambda_k B_k U_k(x, y, z).$$

da cui, come prima, ricaviamo i valori di B_k:

$$B_k = \frac{1}{\lambda_k} \iiint_{\Omega} f_1(x, y, z)\, U_k(x, y, z)\, dx\, dy\, dz.$$

Conoscendo A_k e B_k, conosciamo sia le fasi sia le ampiezze di tutte le oscillazioni proprie.

Abbiamo così mostrato che, sommando oscillazioni proprie, si può ottenere la soluzione più generale del problema con condizioni al contorno omogenee.

Ogni soluzione è dunque composta da oscillazioni proprie, delle quali possiamo calcolare ampiezza e fase conoscendo le condizioni iniziali.

Nello stesso modo si studiano oscillazioni con un numero minore di variabili indipendenti. Consideriamo, per esempio, una corda vibrante fissata a entrambe le estremità. La sua equazione è

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}.$$

Cerchiamo la soluzione per una corda di lunghezza l, fissata agli estremi:

$$u\,|_{x=0} = u\,|_{x=l} = 0.$$

Cerchiamo una famiglia di soluzioni particolari

$$u_k = T_k(t)\, U_k(x).$$

Come prima, otteniamo evidentemente

$$T_k'' U_k = a^2 U_k'' T_k\,,$$

ossia

$$\frac{T_k''}{T_k} = a^2 \frac{U_k''}{U_k} = -\lambda_k^2\,.$$

Da cui

$$T_k = A_k \cos \lambda_k t + B_k \sin \lambda_k t,$$

$$U_k = M_k \cos \frac{\lambda_k}{a} x + N_k \sin \frac{\lambda_k}{a} x.$$

Usiamo le condizioni al contorno per determinare λ_k. Per un valore arbitrario di questo parametro non è possibile soddisfarle entrambe con una soluzione non nulla. Da $U_k|_{x=0} = 0$ ricaviamo $M_k = 0$, quindi $U_k = N_k \sin [(\lambda_k/a)x]$. Ponendo $x = l$, troviamo $\sin (\lambda_k l/a) = 0$. Ciò accade soltanto se $\lambda_k l/a = k\pi$, con k intero. Pertanto

$$\lambda_k = \frac{ak\pi}{l}.$$

La condizione $\int_0^l U_k^2\, dx = 1$ mostra che $N_k = \sqrt{\frac{2}{l}}$. Infine

$$U_k(x) = \sqrt{\frac{2}{l}} \sin \frac{k\pi x}{l}, \quad T_k = A_k \cos \frac{ak\pi t}{l} B_k \sin \frac{ak\pi t}{l}.$$

Nell'espressione di T_k i termini con coseno e seno sono sommati; il segno + è poco leggibile nell'immagine originale.

Le oscillazioni proprie della corda hanno dunque forma sinusoidale, con un numero intero di mezze onde sull'intera corda. Ogni oscillazione ha una propria pulsazione e queste possono essere ordinate in modo crescente:

$$\frac{a\pi}{l}, 2\frac{a\pi}{l}, 3\frac{a\pi}{l}, \cdots, k\frac{a\pi}{l}, \cdots.$$

È noto che a queste pulsazioni corrispondono proprio le frequenze udibili nelle vibrazioni di una corda sonora. La prima è la frequenza fondamentale; le altre sono quelle degli armonici superiori. Le autofunzioni $\sqrt{2/l} \sin (k\pi x/l)$ nell'intervallo $0 \leqslant x \leqslant l$ cambiano segno $k - 1$ volte: infatti $k\pi x/l$ varia da 0 a $k\pi$ e il suo seno cambia segno $k - 1$ volte. I punti in cui le autofunzioni U_k si annullano si chiamano nodi delle oscillazioni.

Se facciamo in modo che la corda rimanga immobile in un punto corrispondente, per esempio, al nodo del primo armonico superiore, il suono fondamentale viene soppresso e si mette in evidenza un suono un'ottava più alto. Questo procedimento, realizzato sfiorando la corda nel nodo, si usa negli strumenti ad arco: violino, viola e violoncello.

Abbiamo esaminato la separazione delle variabili nel problema delle oscillazioni proprie. Il metodo ha però un campo di applicazione molto più vasto, che comprende la conduzione del calore e numerosi altri problemi.

Per l'equazione del calore

$$\Delta T = \frac{\partial T}{\partial t}$$

con la condizione

$$T|_S = 0$$

avremo, come prima,

$$T = \Sigma F_k(t)\, U_k(x, y, z).$$

Qui

$$\frac{F_k'(t)}{F_k(t)} = -\lambda_k^2\,,\ \Delta U_k + \lambda_k^2 U_k = 0.$$

La soluzione si ottiene nella forma

$$T = \sum_{k=1}^{\infty} e^{-\lambda_k^2 t}\, U_k(x, y, z).$$

I coefficienti determinati dalla temperatura iniziale sono qui assorbiti nelle autofunzioni U_k.

Il metodo è stato applicato con grande successo anche ad altre equazioni. Consideriamo, per esempio, l'equazione di Laplace

$$\Delta u = 0$$

nel disco

$$x^2 + y^2 \leqslant 1,$$

e supponiamo di dover costruire una soluzione che soddisfi la condizione

$$u|_{r=1} = f(\vartheta),$$

dove r e ϑ sono le coordinate polari di un punto del piano.

L'equazione di Laplace si trasforma facilmente in coordinate polari e assume la forma

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \vartheta^2} = 0.$$

Cerchiamo una soluzione nella forma

$$u = \sum_{k=1}^{\infty} R_k(r)\, \theta_k(\vartheta).$$

Richiedendo che ogni termine della serie soddisfi separatamente l'equazione, otteniamo

$$\left[R_k''(r) + \frac{1}{r} R_k'(r)\right] \theta_k(\vartheta) + \frac{1}{r^2}\, \theta_k''(\vartheta)\, R_k(r) = 0.$$

Dividendo per $R_k(r)\theta_k(\vartheta)/r^2$, si ha

$$\frac{r^2\left[R_k''(r)+\frac{1}{r}R_k'(r)\right]}{R_k(r)}=-\frac{\theta_k'(\vartheta)}{\theta_k(\vartheta)}.$$

Ponendo ancora

$$\frac{\theta_k''(\vartheta)}{\theta_k(\vartheta)}=-\lambda_k^2,$$

otteniamo

$$r^2\left[R_k''+\frac{1}{r}R_k'\right]-\lambda_k^2R_k=0.$$

È facile vedere che $\theta_k(\vartheta)$ deve essere periodica di periodo 2π. Integrando l'equazione $\theta_k''(\vartheta)+\lambda_k^2\theta_k(\vartheta)=0$,, troviamo

$$\theta_k=a_k\cos\lambda_k\vartheta+b_k\sin\lambda_k\vartheta.$$

Questa funzione ha il periodo richiesto soltanto se λ_k è intero. Ponendo $\lambda_k = k$, otteniamo

$$\theta_k=a_k\cos k\vartheta+b_k\sin k\vartheta.$$

L'equazione per R_k ha, per $k > 0$, la soluzione generale

$$R_k=Ar^k+\frac{B}{r^k}.$$

Conservando soltanto i termini limitati per $r \to 0$, e includendo per $k = 0$ il termine costante, otteniamo la soluzione generale dell'equazione di Laplace nella forma

$$u=a_0+\sum_{k=1}^{\infty}(a_k\cos k\vartheta+b_k\sin k\vartheta)\,r^k.$$

Questo procedimento permette spesso di trovare soluzioni non banali dell'equazione $\Delta U_k+\lambda_k^2U_k=0$ che soddisfino condizioni al contorno omogenee. Quando il problema si riduce alla soluzione di equazioni differenziali ordinarie, si dice che ammette separazione completa delle variabili. Come ha mostrato il matematico sovietico V. V. Stepanov, la separazione completa mediante il metodo di Fourier è possibile soltanto in determinati casi particolari. Il metodo era noto da molto tempo: nelle sue linee essenziali fu usato da Eulero, Bernoulli e d'Alembert. Fourier lo applicò sistematicamente ai problemi della fisica matematica, soprattutto alla conduzione del calore. Tuttavia, come abbiamo detto, spesso esso non è applicabile e bisogna ricorrere ad altri metodi, che ora esamineremo.

Il metodo dei potenziali.

Anche questo metodo si basa sulla sovrapposizione di soluzioni particolari per costruire una soluzione di forma generale. Questa volta, però, come soluzioni fondamentali si usano funzioni che diventano infinite in un punto. Illustriamo il metodo per le equazioni di Laplace e di Poisson.

Sia M_0 un punto dello spazio. Indichiamo con $r(M, M_0)$ la distanza fra M_0 e un punto variabile M. Fissato M_0, la funzione $1/r(M, M_0)$ dipende dal punto M. È facile verificare che è armonica in tutto lo spazio,[5] tranne, naturalmente, nel punto M_0, dove essa e le sue derivate diventano infinite.

La somma di più funzioni di questo tipo

$$\sum_{i=1}^{N} A_i \frac{1}{r(M, M_i)},$$

dove M_1, M_2, ..., M_N sono punti qualsiasi dello spazio, è ancora armonica rispetto a M. Essa presenta singolarità nei punti M_i. Distribuendo questi punti sempre più fittamente in un volume Ω e moltiplicando nello stesso tempo i termini per opportuni coefficienti A_i, possiamo passare al limite e ottenere una nuova funzione

$$U = \lim \sum_{i=1}^{N} \frac{A_i}{r(M, M_i)} = \iiint_{\Omega} \frac{A(M')}{r(M, M')} d\Omega,$$

dove M' percorre l'intero volume Ω. Un integrale di questa forma si chiama potenziale newtoniano. Si può dimostrare, benché qui non lo faremo, che la funzione U così costruita soddisfa l'equazione $\Delta U = -4\pi A$.

Il potenziale newtoniano ha un semplice significato fisico. Per comprenderlo, cominciamo dalla funzione $A_i/r(M, M_i)$.

Le sue derivate parziali rispetto alle coordinate sono

$$A_i \frac{x_i - x}{r^3} = X, \quad A_i \frac{y_i - y}{r^3} = Y, \quad A_i \frac{z_i - z}{r^3} = Z.$$

Poniamo nel punto M_i una massa A_i, che attrae tutti i corpi con una forza diretta verso M_i e inversamente proporzionale al quadrato della distanza. Scomponiamo la forza lungo gli assi coordinati. Se l'intensità della forza agente su una massa puntiforme unitaria è A_i/r^2, i coseni degli angoli che la sua direzione forma con gli assi sono $(x_i - x)/r$, $(y_i - y)/r$, $(z_i - z)/r$. Le componenti della forza esercitata dal centro attrattivo M_i sulla massa unitaria posta in M sono quindi X, Y e Z, ossia le derivate parziali di A_i/r rispetto alle coordinate. Se poniamo masse attrattive nei punti M_1, M_2, ..., M_N, sulla massa unitaria in M agisce la risultante di tutte le loro forze. In altri termini,

$$X = \frac{\partial}{\partial x}\sum \frac{A_i}{r(M, M_i)}, \quad Y = \frac{\partial}{\partial y}\sum \frac{A_i}{r(M, M_i)}, \quad Z = \frac{\partial}{\partial z}\sum \frac{A_i}{r(M, M_i)}.$$

Passando al limite e sostituendo la somma con un integrale, otteniamo

$$\bar{X} = \frac{\partial U}{\partial x}, \quad \bar{Y} = \frac{\partial U}{\partial y}, \quad \bar{Z} = \frac{\partial U}{\partial z}, \quad \text{where} \quad U = \iiint_{\Omega} \frac{A}{r}\, d\Omega.$$

Nella formula: «where» = «dove».

La funzione U, le cui derivate parziali sono le componenti della forza, si chiama qui potenziale della forza. Pertanto $A_i/r(M, M_i)$ è il potenziale dell'attrazione esercitata dal punto M_i; $\Sigma[A_i/r(M, M_i)]$ è quello del gruppo di punti $M_1, M_2, \ldots, M_N$; la funzione $U = \iiint_\Omega (A/r)\, d\Omega$ è il potenziale dell'attrazione esercitata dalle masse distribuite con continuità nel volume Ω.

Invece di distribuire le masse in un volume, possiamo porre i punti $M_1, M_2, \ldots, M_N$ su una superficie S. Aumentandone il numero, otteniamo al limite l'integrale

$$V = \iint_S \frac{A(Q)}{r}\, ds, \qquad (22)$$

dove Q è un punto della superficie S.

È facile vedere che questa funzione è armonica ovunque all'interno e all'esterno di S. Sulla superficie stessa la funzione è continua, come si può dimostrare, mentre le sue derivate prime possono presentare salti finiti.

Anche $\partial(1/r)/\partial x_i$, $\partial(1/r)/\partial y_i$ e $\partial(1/r)/\partial z_i$ sono funzioni armoniche del punto M per M_i fissato. Con esse possiamo formare le somme

$$\sum A_i \frac{\partial \frac{1}{r}}{\partial x_i} + \sum B_i \frac{\partial \frac{1}{r}}{\partial y_i} + \sum C_i \frac{\partial \frac{1}{r}}{\partial z_i},$$

che sono armoniche ovunque, salvo eventualmente nei punti $M_1, M_2, \ldots, M_N$.

Particolare importanza ha l'integrale

$$W = \iint_S \mu(Q) \left[\frac{\partial \frac{1}{r}}{\partial x'} \cos(n, x) + \frac{\partial \frac{1}{r}}{\partial y'} \cos(n, y) + \frac{\partial \frac{1}{r}}{\partial z'} \cos(n, z)\right] ds$$
$$= \iint_S \mu(Q)\, K(Q, M)\, ds, \qquad (23)$$

nel quale x', y' e z' sono le coordinate di un punto variabile Q della superficie S, n è la direzione della normale a S in Q, mentre x, y e z indicano le direzioni degli assi coordinati; r è la distanza fra Q e il punto M nel quale si calcola W.

L'integrale (22) si chiama potenziale di strato semplice e il (23) potenziale di doppio strato.[6] Entrambi rappresentano funzioni armoniche all'interno e all'esterno della superficie *S*.

Molti problemi della teoria delle funzioni armoniche si risolvono mediante i potenziali. Usando il potenziale di doppio strato possiamo costruire, in un dominio dato, una funzione armonica u che assuma i valori assegnati $2\pi\phi(Q)$ sul contorno S. Basta scegliere opportunamente la funzione $\mu(Q)$.

Il problema ricorda quello di determinare i coefficienti della serie

$$\phi = \sum a_k U_k$$

affinché essa rappresenti la funzione a primo membro.

Una notevole proprietà dell'integrale W è che il suo valore limite, quando M tende a Q_0 dall'interno della superficie, ha la forma

$$\lim_{M \to Q_0} W = 2\pi\mu(Q_0) + \iint_S K(Q, Q_0)\, \mu(Q)\, ds.$$

Uguagliando questa espressione alla funzione assegnata $2\pi\phi(Q_0)$, otteniamo l'equazione

$$\mu(Q_0) + \frac{1}{2\pi} \iint_S K(Q, Q_0)\, \mu(Q)\, ds = \phi(Q_0).$$

Questa si chiama equazione integrale di seconda specie. La teoria di tali equazioni è stata sviluppata da molti matematici. Risolvendola con un metodo qualsiasi, otteniamo una soluzione del problema originario.

Analogamente si risolvono altri problemi della teoria delle funzioni armoniche. Scelto un potenziale adatto, se ne determina la densità, cioè la funzione arbitraria che vi compare, in modo da soddisfare tutte le condizioni prescritte.

Dal punto di vista fisico, ciò significa che una funzione armonica nel dominio considerato può essere rappresentata come potenziale di un doppio strato elettrico distribuito su S con densità opportuna.

Costruzione approssimata delle soluzioni; metodo di Galerkin e metodo dei reticoli.

1. Abbiamo discusso due metodi per risolvere le equazioni della fisica matematica: la separazione completa delle variabili e il metodo dei potenziali. Essi furono sviluppati dagli scienziati del XVIII e XIX secolo: Fourier, Poisson, Ostrogradskij, Ljapunov e altri. Nel XX secolo si aggiunsero numerosi nuovi metodi. Ne esamineremo due: il metodo di Galerkin e quello delle differenze finite, o dei reticoli.

Il primo fu proposto dall'accademico B. G. Galerkin per risolvere equazioni della forma

$$\sum\sum\sum\sum A_{ijkl}\frac{\partial^4 U}{\partial x_i\,\partial x_j\,\partial x_k\,\partial x_l}+\sum\sum\sum B_{ijk}\frac{\partial^3 U}{\partial x_i\,\partial x_j\,\partial x_k}$$
$$+\sum\sum C_{ij}\frac{\partial^2 U}{\partial x_i\,\partial x_j}+\sum D_i\frac{\partial U}{\partial x_i}+EU+\lambda U=0,$$

contenenti un parametro incognito λ, dove gli indici i, j, k e l assumono indipendentemente i valori 1, 2 e 3. Queste equazioni si ricavano da equazioni contenenti la variabile indipendente t mediante la separazione delle variabili, così come dall'equazione delle onde

$$\Delta u=\frac{\partial^2 u}{\partial t^2}$$

si ottiene $\Delta U + \lambda^2 U = 0$. Il problema consiste nel trovare i valori di λ per i quali il problema al contorno omogeneo ammette una soluzione non nulla e poi nel costruirla.

L'idea del metodo di Galerkin è la seguente. Si cerca la funzione incognita nella forma approssimata

$$U\approx\sum_{m=1}^{N}a_m\omega_m(x_1,x_2,x_3),$$

dove le $\omega_m(x_1, x_2, x_3)$ sono funzioni scelte che soddisfano le condizioni al contorno.

Sostituendo questa espressione nel primo membro dell'equazione, si ottiene l'equazione approssimata

$$\sum_{m=1}^{N}a_m\left[\sum\sum\sum\sum A_{ijkl}\frac{\partial^4\omega_m}{\partial x_i\,\partial x_j\,\partial x_k\,\partial x_l}+\sum\sum\sum B_{ijk}\frac{\partial^3\omega_m}{\partial x_i\,\partial x_j\,\partial x_k}\right.$$
$$\left.+\sum\sum C_{ij}\frac{\partial^2\omega_m}{\partial x_i\,\partial x_j}+\sum D_i\frac{\partial\omega_m}{\partial x_i}+E\omega_m\right]+\lambda\sum_{m=1}^{N}a_m\omega_m\approx 0.$$

Indichiamo per brevità con $L\omega_m$ l'espressione fra parentesi e scriviamo

$$\sum a_m L\omega_m+\lambda\sum a_m\omega_m\approx 0.$$

Moltiplichiamo entrambi i membri per ω_n e integriamo sul dominio Ω in cui cerchiamo la soluzione. Otteniamo

$$\iiint_\Omega\sum a_m\omega_n L\omega_m\,d\Omega+\lambda\iiint_\Omega\sum a_m\omega_m\omega_n\,d\Omega\approx 0,$$

che si può riscrivere nella forma

$$\sum_{m=1}^{N}a_m\iiint_\Omega\omega_n L\omega_m\,d\Omega+\lambda\sum_{m=1}^{N}a_m\iiint_\Omega\omega_m\omega_n\,d\Omega\approx 0.$$

Imponendo che queste equazioni siano soddisfatte esattamente, otteniamo un sistema di equazioni algebriche di primo grado nei coefficienti incogniti a_m. Il numero delle equazioni uguaglia quello delle incognite; essendo il sistema omogeneo, esso ha una soluzione non nulla soltanto se il determinante si annulla. Sviluppando il determinante, ricaviamo un'equazione di grado N nel numero incognito λ.

Trovato un valore di λ e sostituitolo nel sistema, risolviamo quest'ultimo e otteniamo un'espressione approssimata della funzione U.

Il metodo di Galerkin non è adatto soltanto alle equazioni del quarto ordine: si applica a equazioni di ordini e tipi diversi.

2. L'ultimo metodo che esamineremo si chiama metodo delle differenze finite, o metodo dei reticoli.

La derivata della funzione u rispetto a x è definita come limite del rapporto

$$\frac{u(x+\Delta x)-u(x)}{\Delta x}.$$

Questo rapporto può a sua volta essere rappresentato nella forma

$$\frac{1}{\Delta x}\int_x^{x+\Delta x}\frac{\partial u}{\partial x_1}\,dx_1,$$

e, per il noto teorema del valor medio (cfr. capitolo II, §8),

$$\frac{u(x+\Delta x)-u(x)}{\Delta x}=\frac{\partial u}{\partial x}\Big|_{x=\xi},$$

dove ξ è un punto dell'intervallo

$$x<\xi<x+\Delta x.$$

Anche tutte le derivate seconde, sia quelle miste sia quelle rispetto a una sola variabile, possono essere rappresentate approssimativamente mediante rapporti di differenze. Per esempio, il rapporto

$$\frac{u(x+\Delta x)-2u(x)+u(x-\Delta x)}{(\Delta x)^2}$$

si scrive nella forma

$$\frac{1}{\Delta x}\left[\frac{u(x+\Delta x)-u(x)}{\Delta x}-\frac{u(x)-u(x-\Delta x)}{\Delta x}\right]$$
$$=\frac{1}{\Delta x}\left\{\left[\frac{u(x_1+\Delta x)-u(x_1)}{\Delta x}\right]\Big|_{x_1=x-\Delta x}^{x_1=x}\right\}.$$

Per il teorema del valor medio, il rapporto incrementale della funzione

$$\phi(x_1) = \frac{u(x_1 + \Delta x) - u(x_1)}{\Delta x}$$

può essere sostituito da un valore della sua derivata. Di conseguenza

$$\frac{\phi(x_1) - \phi(x_1 - \Delta x)}{\Delta x} = \phi'(\xi),$$

dove ξ è un valore intermedio nell'intervallo

$$x - \Delta x < \xi < x.$$

Dunque

$$\left(\frac{1}{\Delta x}\right)^2 [u(x + \Delta x) - 2u(x) + u(x - \Delta x)] = \frac{1}{\Delta x} [\phi(x) - \phi(x - \Delta x)] = \phi'(\xi).$$

D'altra parte

$$\phi(\xi) = \frac{u(\xi + \Delta x) - u(\xi)}{\Delta x},$$

e quindi

$$\phi'(\xi) = \frac{u'(\xi + \Delta x) - u'(\xi)}{\Delta x}.$$

Applicando ancora il teorema del valor medio, vediamo che

$$\phi'(\xi) = u''(\eta),$$

dove

$$\xi < \eta < \xi + \Delta x.$$

Di conseguenza,

$$\left(\frac{1}{\Delta x}\right)^2 [u(x + \Delta x) - 2u(x) + u(x - \Delta x)] = u''(\eta),$$

dove $x - \Delta x < \eta < x + \Delta x$.

Se $u''(x)$ è continua e Δx è abbastanza piccolo, $u''(\eta)$ differisce di poco da $u''(x)$. La derivata seconda è dunque arbitrariamente vicina al rapporto di differenze considerato. Allo stesso modo si mostra, per esempio, che la derivata seconda mista

$$\frac{\partial^2 u}{\partial x\,\partial y}$$

può essere rappresentata approssimativamente dalla formula

$$\frac{\partial^2 u}{\partial x\,\partial y} = \frac{1}{\Delta x\,\Delta y}[u(x+\Delta x, y+\Delta y) - u(x+\Delta x, y) - u(x, y+\Delta y) + u(x, y)].$$

Torniamo ora alla nostra equazione alle derivate parziali.

Per fissare le idee, consideriamo l'equazione di Laplace in due variabili indipendenti

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Supponiamo inoltre che la funzione incognita u sia assegnata sul contorno S del dominio Ω. Approssimativamente abbiamo

$$\frac{\partial^2 u}{\partial x^2} = \frac{u(x+\Delta x, y) - 2u(x, y) + u(x-\Delta x, y)}{(\Delta x)^2},$$
$$\frac{\partial^2 u}{\partial y^2} = \frac{u(x, y+\Delta y) - 2u(x, y) + u(x, y-\Delta y)}{(\Delta y)^2}$$

Ponendo $\Delta x = \Delta y = h$, otteniamo

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{1}{h^2}[u(x+h, y) + u(x, y+h) + u(x-h, y) + u(x, y-h) - 4u(x, y)].$$

Ricopriamo il dominio Ω con un reticolo quadrato i cui vertici hanno coordinate multiple di h (figura 4). Sostituiamo il dominio con il poligono costituito dai quadrati del reticolo contenuti in Ω; il contorno viene così approssimato da una spezzata. Su questa spezzata attribuiamo alla funzione i valori corrispondenti ai dati assegnati su S. L'equazione di Laplace viene approssimata dall'equazione

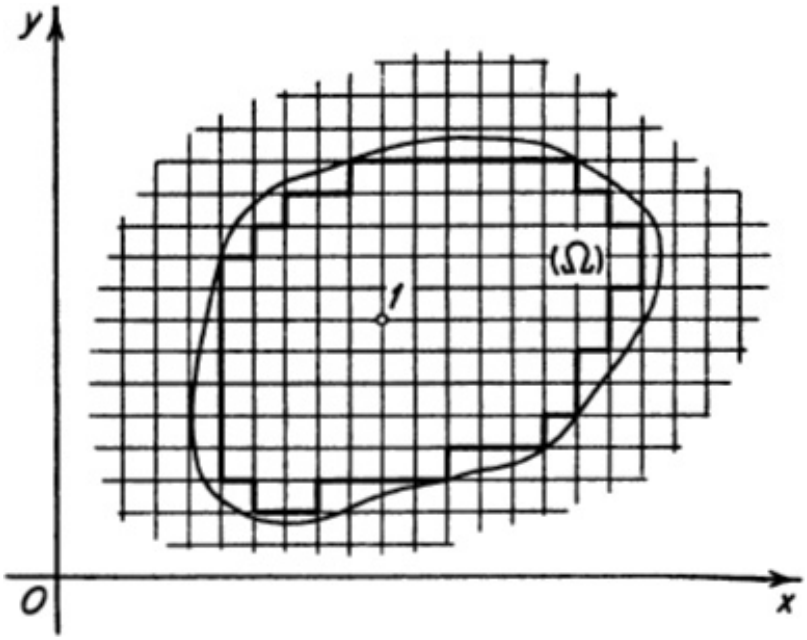


Figura 4.

$$u(x+h, y) + u(x, y+h) + u(x-h, y) + u(x, y-h) - 4u(x, y) = 0$$

in tutti i nodi interni. Essa si può riscrivere nella forma

$$u(x, y) = \tfrac{1}{4}\,[u(x+h, y) + u(x, y+h) + u(x-h, y) + u(x, y-h)].$$

Il valore di u in ogni nodo del reticolo, per esempio nel punto 1 della figura 4, è dunque la media aritmetica dei valori nei quattro nodi adiacenti.

Supponiamo che all'interno del poligono vi siano N nodi. A ciascuno corrisponde un'equazione. Otteniamo così un sistema di N equazioni algebriche in N incognite, la cui soluzione fornisce valori approssimati della funzione u nel dominio Ω.

Si può dimostrare che, per l'equazione di Laplace, in questo modo la soluzione si determina con la precisione desiderata.

Il metodo delle differenze finite riduce quindi il problema alla soluzione di un sistema di N equazioni in N incognite, che sono i valori della funzione cercata nei nodi di un reticolo.

Il metodo si applica anche ad altri problemi della fisica matematica, ad altre equazioni differenziali e alle equazioni integrali. In molti casi, però, la sua applicazione presenta difficoltà.

Può accadere che il sistema di N equazioni costruito mediante il reticolo non abbia soluzione, oppure fornisca un risultato assai lontano da quello esatto. Ciò può verificarsi quando gli errori si accumulano durante la soluzione: quanto più piccole scegliamo le maglie, tanto più numerose sono le equazioni e tanto maggiore può diventare l'errore accumulato.

Nell'esempio precedente dell'equazione di Laplace non si presenta questo tipo di instabilità. Con opportuni procedimenti di approssimazioni successive, gli errori dell'iterazione

diminuiscono progressivamente. Per l'equazione del calore e per quella delle onde è invece essenziale scegliere correttamente il reticolo: si possono ottenere risultati buoni oppure cattivi.

Negli schemi espliciti qui considerati, una volta scelto il passo temporale, non si deve scegliere un reticolo spaziale troppo fine rispetto a esso. Altrimenti si ottiene un sistema inadeguato, la cui soluzione oscilla rapidamente con grandi ampiezze e si allontana molto da quella esatta.

La varietà dei risultati possibili si vede bene in un semplice esempio numerico. Consideriamo l'equazione

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

che descrive la conduzione del calore quando la temperatura non dipende da y né da z. Scegliamo un passo k per il tempo e un passo h per la coordinata x:

$$\frac{\partial u}{\partial t} \approx \frac{u(t + k, x) - u(t, x)}{k},$$
$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u(t, x + h) - 2u(t, x) + u(t, x - h)}{h^2}.$$

L'equazione si scrive allora approssimativamente nella forma

$$u(t + k, x) = \frac{k}{h^2} u(t, x + h) + \left(1 - 2\frac{k}{h^2}\right) u(t, x) + \frac{k}{h^2} u(t, x - h).$$

Se a un certo istante del reticolo conosciamo u nei punti $x - h$, x e $x + h$, è facile determinarne il valore in x all'istante successivo $t + k$. Supponiamo già fissata la costante k, cioè il passo temporale. Consideriamo due scelte di h: nel primo caso $h^2 = k$, nel secondo $h^2 = 2k$. Risolviamo con il metodo dei reticoli il seguente problema.

All'istante iniziale $u = 0$ per tutti i valori negativi di x e $u = 1$ per tutti i valori non negativi. Scrivendo su ogni riga i valori dell'incognita a un dato istante, otteniamo le due tabelle seguenti.

Tabella 1

t \ x	$-5h$	$-4h$	$-3h$	$-2h$	$-h$	0	h	$2h$	$3h$	$4h$	$5h$
0	0	0	0	0	0	1	1	1	1	1	1
k	0	0	0	0	1	0	1	1	1	1	1
$2k$	0	0	0	1	−1	2	0	1	1	1	1
$3k$	0	0	1	−2	4	−3	3	0	1	1	1
$4k$	0	1	−3	7	−9	10	−6	4	0	1	1
$5k$	1	−4	11	−19	26	−25	20	−10	5	0	1

Tabella 2

t \ x	$-5h$	$-4h$	$-3h$	$-2h$	$-h$	0	h	$2h$	$3h$	$4h$	$5h$
0	0	0	0	0	0	1	1	1	1	1	1
k	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	1	1
$2k$	0	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{4}$	1	1	1	1
$3k$	0	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{7}{8}$	$\frac{7}{8}$	1	1	1
$4k$	0	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{11}{16}$	$\frac{11}{16}$	$\frac{15}{16}$	$\frac{15}{16}$	1	1
$5k$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{13}{16}$	$\frac{13}{16}$	$\frac{31}{32}$	$\frac{31}{32}$	1

Nella tabella 2, a ogni istante fissato, i valori variano gradualmente da un punto all'altro. Essa fornisce una buona approssimazione della soluzione dell'equazione del calore. Nella tabella 1, invece, benché la suddivisione più fine dell'intervallo spaziale sembri dover aumentare la precisione, i valori di u oscillano rapidamente fra valori positivi e negativi e raggiungono grandezze assai superiori a quelle iniziali. È evidente che essi sono enormemente lontani dai valori della soluzione esatta.

Questi esempi mostrano che, per ottenere risultati sufficientemente precisi e affidabili con il metodo dei reticoli, occorre grande attenzione nella scelta dei passi e uno studio preliminare che giustifichi l'applicazione del metodo.

Le soluzioni ottenute mediante le equazioni della fisica matematica, per questi e altri problemi delle scienze naturali, descrivono matematicamente lo svolgimento o il carattere previsto dei fenomeni fisici rappresentati dalle equazioni.

Nella costruzione di un modello mediante tali equazioni siamo costretti, per astrazione, a trascurare molti aspetti dei fenomeni, a scartarne alcuni come non essenziali e a sceglierne altri come fondamentali. I risultati ottenuti non sono quindi assolutamente veri rispetto alla realtà: lo sono soltanto per lo schema o modello considerato. Devono sempre essere confrontati con l'esperienza, per accertare che il modello sia abbastanza vicino al fenomeno e lo rappresenti con precisione sufficiente.

Il criterio ultimo della verità dei risultati è dunque l'esperienza pratica. In definitiva vi è questo solo criterio, benché l'esperienza possa essere interpretata correttamente soltanto alla luce di una teoria profonda e sviluppata.

Considerando la corda vibrante di uno strumento musicale, possiamo capire come essa produca i suoni soltanto conoscendo le leggi di sovrapposizione delle oscillazioni proprie. I rapporti fra le frequenze si comprendono studiando come esse dipendano dal materiale, dalla tensione della corda e dal modo in cui sono fissate le estremità. La teoria non offre soltanto un mezzo per calcolare le quantità numeriche desiderate: indica anche quali siano fondamentali, come si svolga il processo fisico e che cosa occorra osservare.

La fisica matematica, dunque, non è soltanto nata dalle esigenze della pratica: a sua volta ha agito sulla pratica, indicandole nuove vie di progresso.

Essa è strettamente collegata ad altri rami dell'analisi matematica, ma non possiamo discutere qui questi legami, che ci porterebbero troppo lontano.

§6. Soluzioni generalizzate

L'insieme dei problemi nei quali un processo fisico è descritto da funzioni continue e derivabili che soddisfano equazioni differenziali può essere ampliato in modo sostanziale introducendo soluzioni discontinue di queste equazioni.

In molti casi è chiaro fin dall'inizio che il problema non può avere soluzioni con derivate seconde continue. Dal punto di vista della formulazione classica esposta nei paragrafi precedenti, esso è quindi privo di soluzione. Eppure il processo fisico corrispondente si svolge realmente, anche se non troviamo funzioni che lo descrivano nella classe prestabilita delle funzioni regolari. Consideriamo alcuni semplici esempi.

1. Se una corda è composta da due tratti di densità diversa, nell'equazione

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \qquad (24)$$

il coefficiente assume due valori costanti diversi nei due tratti. Di conseguenza, per dati generali, l'equazione (24) non ammette una soluzione classica, cioè con derivate seconde continue in tutto il dominio.

2. Il coefficiente a sia costante, ma la posizione iniziale della corda sia una spezzata descritta da $u|_{t=0} = \phi(x)$. Nei vertici la funzione ϕ non ha evidentemente derivata prima. Si può mostrare che non esiste una soluzione classica della (24) che soddisfi le condizioni iniziali

$$u|_{t=0} = \phi(x), \quad u_t|_{t=0} = 0$$

Qui e nel seguito u_t indica $\partial u/\partial t$.

3. Se si dà un colpo brusco a un piccolo tratto della corda, le oscillazioni risultanti sono descritte dall'equazione

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t),$$

dove $f(x, t)$ rappresenta l'azione esercitata: è una funzione discontinua, diversa da zero soltanto sul piccolo tratto e durante un breve intervallo di tempo. Anche questa equazione, come è facile verificare, non può avere una soluzione con derivate seconde continue che la soddisfi ovunque.

Questi esempi mostrano che la richiesta di derivate continue limita fortemente l'insieme dei problemi risolvibili. Per ampliarlo si cominciò ad ammettere discontinuità di prima specie nelle derivate di ordine massimo delle funzioni cercate, richiedendo che esse soddisfacessero l'equazione fuori dai punti di discontinuità. Risulta che le soluzioni di equazioni del tipo $\Delta u = 0$ o $\partial u/\partial t - \Delta u = 0$ non possono presentare, all'interno del dominio, simili discontinuità, dette deboli. Le soluzioni dell'equazione delle onde possono invece avere discontinuità deboli nelle variabili x, y, z e t, ma soltanto su superfici di forma particolare, chiamate superfici caratteristiche. Se consideriamo una soluzione $u(x, y, z, t)$ come un campo scalare nello spazio all'istante $t = t_1$, le superfici di discontinuità delle sue derivate seconde si propagano nello spazio con velocità pari alla radice quadrata del coefficiente del laplaciano nell'equazione delle onde.

Il secondo esempio della corda mostra che bisogna considerare anche soluzioni con derivate prime discontinue; per le onde sonore e luminose occorre persino considerare soluzioni che siano esse stesse discontinue.

La prima questione è stabilire quali funzioni discontinue possano essere ammesse come soluzioni fisicamente significative dell'equazione o del problema corrispondente. Potremmo, per esempio, essere tentati di considerare una funzione arbitraria costante a tratti come «un'unica soluzione» dell'equazione di Laplace o di quella delle onde, poiché soddisfa l'equazione fuori dagli insiemi di discontinuità.

Per chiarire la questione bisogna anzitutto assicurare che, nella classe più ampia cui appartengono le soluzioni ammissibili, valga ancora un teorema di unicità. È evidente che, ammettendo funzioni regolari a tratti del tutto arbitrarie, questa richiesta non sarebbe soddisfatta.

Storicamente, il primo criterio fu richiedere che le funzioni ammissibili fossero limiti, in un senso opportuno, di soluzioni classiche della stessa equazione. Nell'esempio 2, una soluzione

della (24) corrispondente a $\phi(x)$, non derivabile in un punto angoloso, si trova così come limite uniforme di soluzioni classiche $u_n(x, t)$ con condizioni iniziali $u_n|_{t=0} = \phi_n(x)$ e $\partial u_n/\partial t|_{t=0} = 0$, dove le ϕ_n hanno derivate seconde continue e convergono uniformemente a ϕ per $n \to \infty$.

Nel seguito adotteremo invece questo criterio: una soluzione ammissibile u deve soddisfare, al posto dell'equazione $Lu = f$, un'identità integrale contenente una funzione arbitraria Φ.

L'identità si ottiene così. Moltiplichiamo entrambi i membri di $Lu = f$ per una funzione arbitraria Φ che abbia derivate continue, rispetto a tutte le variabili, fino all'ordine dell'equazione e supporto compatto all'interno del dominio limitato D in cui l'equazione è definita. Integriamo su D e trasformiamo mediante integrazione per parti fino a eliminare tutte le derivate di u. Otteniamo l'identità desiderata. Per l'equazione (24), per esempio, essa ha la forma

$$\iint_D u \left[\frac{\partial^2 \Phi}{\partial t^2} - \frac{\partial^2 (a^2 \Phi)}{\partial x^2} \right] dx\, dt = 0.$$

S. L. Sobolev ha mostrato che, per le equazioni a coefficienti costanti, questi due criteri di selezione delle soluzioni ammissibili, oggi abitualmente dette generalizzate, sono equivalenti nel senso appropriato. Per le equazioni a coefficienti variabili, invece, il primo può risultare inapplicabile, poiché in generale possono mancare soluzioni classiche del problema (cfr. esempio 1). Il secondo permette di definire soluzioni generalizzate sotto ipotesi molto ampie sulla regolarità dei coefficienti. A prima vista, però, esso può sembrare eccessivamente formale e puramente matematico: non indica direttamente come formulare i problemi in modo analogo a quelli classici.

Presenteremo una modifica che appare più aderente al significato fisico, essendo direttamente collegata al noto principio di Hamilton.

Come è noto, l'analisi dei procedimenti con cui si ricavano le equazioni della fisica matematica condusse, nella prima metà del XIX secolo, alla scoperta del principio di Hamilton. Esso consentì di ottenere in modo unitario un'ampia classe di equazioni della fisica matematica. Lo illustreremo con il problema, considerato nel §3, delle oscillazioni di una corda di lunghezza finita con estremità fisse.

Costruiamo anzitutto la cosiddetta funzione di Lagrange $L(t)$ della corda, cioè la differenza fra energia cinetica ed energia potenziale. Da quanto detto nel §3 segue che

$$L(t) = \int_0^l \left(\frac{1}{2} \rho u_t^2 - \frac{T}{2} u_x^2 \right) dx.$$

Secondo il principio di Hamilton, l'integrale

$$S = \int_{t_1}^{t_2} L(t)\, dt$$

è stazionario per la funzione $u(x, t)$ che corrisponde al moto effettivo della corda, rispetto alle altre funzioni $v(x, t)$ nulle per $x = 0$ e $x = l$ e coincidenti con $u(x, t_1)$ e $u(x, t_2)$ ai tempi t_1 e t_2. Questi tempi sono fissati arbitrariamente e gli integrali S delle funzioni ammissibili devono essere finiti. Il principio richiede che la cosiddetta variazione prima di S (cfr. capitolo VIII) sia nulla, cioè

$$\delta S = \int_{t_1}^{t_2} \int_0^l (\rho u_t \Phi_t - T u_x \Phi_x)\, dx\, dt = 0, \qquad (25)$$

dove $\Phi(x, t)$ è una funzione arbitraria derivabile rispetto a x e t e nulla sui lati del rettangolo $0 \leqslant x \leqslant l$, $t_1 \leqslant t \leqslant t_2$.

La (25) è dunque la condizione che deve soddisfare la funzione cercata $u(x, t)$. Se sappiamo che essa possiede derivate seconde, possiamo trasformare questa condizione. Integrando per parti e applicando il lemma fondamentale del calcolo delle variazioni, troviamo che u deve soddisfare l'equazione

$$\frac{\partial}{\partial t}\left(\rho \frac{\partial u}{\partial t}\right) - \frac{\partial}{\partial x}\left(T \frac{\partial u}{\partial x}\right) = 0, \qquad (26)$$

che coincide con la (24) quando ρ e T sono costanti e $T/\rho = a^2$.

È facile vedere che ogni soluzione classica $u(x, t)$ della (26) soddisfa l'identità (25) per tutte le funzioni Φ considerate. Il viceversa è falso, perché u può non avere derivate seconde. Sostituire l'equazione (26) con l'identità (25) amplia dunque l'insieme dei problemi risolvibili.

Per determinare una particolare oscillazione della corda, alle condizioni al contorno

$$u(0, t) = u(l, t) = 0, \qquad (27)$$

dobbiamo aggiungere le condizioni iniziali

$$\begin{aligned} u(x, 0) &= \phi_0(x), \\ u_t(x, 0) &= \phi_1(x). \end{aligned} \qquad (28)$$

Se cerchiamo una soluzione fra le funzioni con derivate prime continue, le condizioni (27) e (28) possono essere imposte separatamente dalla (25). Se ammettiamo invece soluzioni meno regolari, esse perdono significato nella forma data e devono essere incorporate, in parte o interamente, nell'identità integrale.

Per esempio, $u(x, t)$ sia continua per $0 \leqslant x \leqslant l, 0 \leqslant t \leqslant T$, ma abbia derivate prime discontinue. La seconda condizione (28) non è allora necessariamente interpretabile come condizione puntuale al limite. Il problema può essere formulato così: trovare una funzione continua u che soddisfi la (27) e la prima delle (28) e per la quale l'uguaglianza

$$\int_0^T \int_0^l (\rho u_t \Phi_t - T u_x \Phi_x)\, dx\, dt + \int_0^l \phi_1 \Phi(x, 0)\, dx = 0 \qquad (29)$$

Nelle formule (29) e nel funzionale successivo, con φ_1 definita come velocità iniziale dalla (28), il termine iniziale va moltiplicato per ρ. La forma riprodotta equivale a usare la quantità di moto iniziale, oppure a normalizzare ρ = 1.

valga identicamente per tutte le funzioni continue $\Phi(x, t)$ nulle per $x = 0$, $x = l$ e $t = T$. Le funzioni u e Φ devono avere derivate prime i cui quadrati siano integrabili nel senso di Lebesgue sul rettangolo $0 \leqslant x \leqslant l$, $0 \leqslant t \leqslant T$. Per u, quest'ultima richiesta significa che la media temporale dell'energia totale della corda

$$\frac{1}{2T}\int_0^T\int_0^l (\rho u_t^2 + T u_x^2)\,dx\,dt$$

deve essere finita. Una simile restrizione su u, e quindi sulle sue possibili variazioni Φ, deriva naturalmente dal principio di Hamilton.

L'identità (29) esprime precisamente l'annullarsi della variazione prima del funzionale

$$\tilde{S} = \int_0^T\int_0^l \left(\frac{\rho}{2}u_t^2 - \frac{T}{2}u_x^2\right)dx\,dt + \int_0^l \phi_1 u\,|_{t=0}\,dx$$

Il problema della corda fissata agli estremi può quindi essere formulato come ricerca di un punto stazionario del funzionale $\tilde{S}$ fra le funzioni continue $v(x, t)$ che soddisfano la (27) e coincidono con $u(x, T)$ per $t = T$. La funzione cercata deve inoltre soddisfare la prima delle condizioni (28).

Questa modifica del principio di Hamilton permette non soltanto di ampliare la classe delle soluzioni ammissibili della (24), ma anche di formulare per esse un problema al contorno ben definito.

Il fatto che queste soluzioni generalizzate, o alcune loro derivate, non siano definite in tutti i punti dello spazio non contraddice l'esperienza. Lo sottolineò ripetutamente N. M. Gjunter, le cui ricerche contribuirono in modo decisivo ad affermare una nuova concezione della soluzione di un'equazione della fisica matematica.

Per esempio, per determinare il moto di un liquido in un canale, nella formulazione classica dobbiamo calcolare il vettore velocità e la pressione in ogni punto. Nella pratica, però, non misuriamo la pressione in un punto, ma la sua azione su una certa area; né la velocità in un singolo punto, ma la quantità di liquido che attraversa un'area nell'unità di tempo. La definizione di soluzione generalizzata mira quindi, in sostanza, proprio al calcolo delle grandezze dotate di significato fisico diretto.

Per rendere risolvibile un maggior numero di problemi, dobbiamo cercare le soluzioni nella classe di funzioni più ampia possibile per la quale valgano ancora teoremi di unicità. Spesso questa classe è suggerita dalla natura fisica del problema. In meccanica quantistica, per esempio, il significato fisico non è legato al valore puntuale della funzione di stato $\psi(x)$, soluzione dell'equazione di Schrödinger, ma a quantità integrali come $a_\nu = \int_E \psi(x)\psi_\nu(x)dx$, dove le ψ_ν sono funzioni per le quali $\int_E \psi_\nu^2\,dx < \infty$. Si cerca quindi ψ fra le funzioni a quadrato integrabile, anziché limitarsi a quelle con derivate seconde continue. Per i problemi

dell'elettrodinamica quantistica, all'epoca dell'opera rimaneva aperta la scelta delle classi di funzioni nelle quali cercare le soluzioni delle equazioni considerate.

I progressi della fisica matematica nei trent'anni precedenti la stesura dell'opera sono stati strettamente legati a questa nuova formulazione dei problemi e alla creazione degli strumenti matematici necessari per risolverli. Fra questi ha un ruolo centrale il cosiddetto teorema di immersione di S. L. Sobolev.

Metodi particolarmente convenienti per trovare soluzioni generalizzate in queste classi sono le differenze finite, i metodi diretti del calcolo delle variazioni (Ritz e Trefftz), il metodo di Galerkin e i metodi degli operatori funzionali. Questi ultimi si fondano essenzialmente sullo studio delle trasformazioni generate dai problemi. Nel §5 abbiamo già parlato delle differenze finite e del metodo di Galerkin. Qui illustreremo le idee fondamentali dei metodi diretti del calcolo delle variazioni.

Consideriamo il problema di determinare la posizione di una membrana uniformemente tesa con contorno fissato. Per il principio del minimo dell'energia potenziale in uno stato di equilibrio stabile, la funzione $u(x, y)$ deve rendere minimo l'integrale

$$J(u) = \iint_D (u_x^2 + u_y^2)\, dx\, dy$$

rispetto a tutte le altre funzioni $v(x, y)$ con derivate prime continue che soddisfano la stessa condizione al contorno $v|_S = \phi$. Con opportune ipotesi su ϕ e su S, si dimostra che il minimo esiste ed è raggiunto da una funzione armonica. La funzione cercata u risolve quindi il problema di Dirichlet $\Delta u = 0$, $u|_S = \phi$. Vale anche il viceversa: la soluzione del problema di Dirichlet minimizza J fra tutte le funzioni ammissibili con quel dato al contorno.

L'esistenza della funzione u che minimizza J e il suo calcolo con precisione arbitraria possono essere ottenuti, per esempio, con il procedimento seguente, detto metodo di Ritz. Scegliamo una famiglia infinita di funzioni con derivate seconde continue $\{v_n(x, y)\}$, $n = 0, 1, 2, \ldots$, nulle sul contorno per $n > 0$ e uguali a ϕ per $n = 0$. Consideriamo J sulle funzioni della forma

$$v = \sum_{k=1}^{n} C_k v_k + v_0\,,$$

dove n è fissato e i C_k sono numeri arbitrari. Allora $J(v)$ è un polinomio di secondo grado nelle n variabili indipendenti $C_1, C_2, \ldots, C_n$. Determiniamo i coefficienti imponendo che il polinomio assuma il minimo. Otteniamo un sistema di n equazioni algebriche lineari in n incognite, il cui determinante è non nullo se le funzioni di prova sono scelte linearmente indipendenti. I coefficienti sono quindi determinati univocamente. Indichiamo la funzione corrispondente con $v^n(x, y)$. Si può dimostrare che, se il sistema $\{v_n\}$ soddisfa un'opportuna condizione di «completezza», le funzioni v^n convergono, per $n \to \infty$, alla soluzione cercata.

In conclusione, osserviamo che in questo capitolo abbiamo descritto soltanto i più semplici problemi lineari della meccanica, lasciando da parte molte altre questioni, ancora lontane

dall'essere completamente risolte all'epoca dell'opera, connesse con equazioni alle derivate parziali più generali.

Letture consigliate

H. Bateman, Partial differential equations of mathematical physics, Dover, New York, 1944.

R. Courant e D. Hilbert, Methods of mathematical physics. II, Partial differential equations, Interscience, New York, 1962.

G. F. D. Duff, Partial differential equations, University of Toronto Press, 1956.

G. E. Forsythe e W. R. Wasow, Finite-difference methods for partial differential equations, Wiley, New York, 1960.

L. Hopf, Introduction to the differential equations of physics, Dover, New York, 1948.

I. G. Petrovskii, Lectures on partial differential equations, Interscience, New York, 1954.

H. Sagan, Boundary and eigenvalue problems in mathematical physics, Wiley, New York, 1961.

I. N. Sneddon, Elements of partial differential equations, McGraw-Hill, New York, 1957.

A. Sommerfeld, Partial differential equations in physics, Academic Press, New York, 1949.

A. G. Webster, Partial differential equations of mathematical physics, Dover, New York, 1955.

Note al capitolo VI

[1] Il simbolo Δu abbrevia l'espressione $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 + \partial^2 u/\partial z^2$ e si chiama laplaciano della funzione u.

[2] Qui e nel seguito «funzione arbitraria» significa che alla funzione non sono imposte condizioni particolari, salvo determinate richieste di regolarità.

[3] Se i secondi membri delle condizioni (13) e (14) sono nulli, queste condizioni si chiamano «omogenee».

[4] Se a uno stesso valore di λ corrispondono più funzioni U essenzialmente diverse, cioè linearmente indipendenti, quel valore viene ripetuto un numero corrispondente di volte nella successione dei valori propri λ_k. L'ortogonalità fra funzioni corrispondenti allo stesso valore può essere assicurata scegliendole opportunamente.

[5] Cioè la funzione soddisfa l'equazione di Laplace.

[6] I nomi di questi potenziali sono legati al seguente fatto fisico. Supponiamo di distribuire cariche elettriche sulla superficie S. Esse generano nello spazio un campo elettrico il cui potenziale è rappresentato dall'integrale (22), detto perciò potenziale di strato semplice.

Supponiamo ora che S sia una sottile pellicola isolante. Su un lato distribuiamo, secondo una certa legge, cariche di un segno, per esempio positive; sull'altro distribuiamo, secondo la stessa legge, cariche di segno opposto. Anche i due strati generano un campo elettrico. Nel limite di un doppio strato di dipoli, il suo potenziale è rappresentato dall'integrale (23).

Nota editoriale

La traduzione conserva le sei sezioni, le quattro figure, le due tabelle, le formule e le sei note del capitolo. Sono corretti refusi di trascrizione in indici, lettere greche, esponenti e bibliografia; il richiamo alla nota 3, poco leggibile nella formula originale, è riportato nel testo. Nel bilancio di massa del §2 la colonnina che attraversa la superficie nel tempo dt è quella compresa inizialmente fra ds_2 e ds, non fra ds_2 e ds_1. Le equazioni dei fluidi sono riferite alle ipotesi ideali usate nella derivazione; nel caso incomprimibile considerato la densità è costante.

Nel §3 i principi di massimo sono formulati in modo da includere correttamente le funzioni costanti. Si intendono domini limitati, dati e contorni sufficientemente regolari e, per le soluzioni classiche, condizioni di compatibilità. Per il problema di Neumann stazionario occorre la compatibilità fra sorgenti e flusso totale: quando esiste, la soluzione è determinata soltanto a meno di una costante. Anche le condizioni miste richiedono ipotesi appropriate sui coefficienti. L'argomento energetico della corda richiede condizioni senza lavoro a entrambe le estremità; la prova stabilisce l'unicità, non da sola l'esistenza o l'intera buona posizione del problema.

Nel §4 l'annullamento per interferenza riguarda lo spostamento, non l'energia. La singola onda sferica contiene il fattore 1/r ed è considerata per $r > 0$; una soluzione regolare nell'origine richiede opportune condizioni o una combinazione di onde convergenti e divergenti. Nel §5 λ_k è una pulsazione (la frequenza in cicli per unità di tempo è $\lambda_k/(2\pi)$); l'ortogonalità annulla i termini con $k \neq j$. Per la soluzione radiale di Laplace, il caso $k = 0$ va trattato separatamente e fornisce il termine costante fra le soluzioni limitate. Il potenziale gravitazionale segue la convenzione di segno dell'originale: il suo gradiente dà la forza, con la

costante gravitazionale assorbita nelle unità. Le restrizioni sui passi del reticolo riguardano gli schemi espliciti illustrati, non ogni metodo alle differenze finite.

Nel §6 il principio di Hamilton è espresso correttamente come stazionarietà dell'azione, non come minimo generale: la variazione prima nulla non garantisce un minimo. Diverso è il principio di minimo dell'energia potenziale per la membrana in equilibrio. Le funzioni di prova sono intese con supporto interno appropriato; le derivate delle soluzioni generalizzate vanno interpretate nel senso debole richiesto. Nella scrittura quantistica si conserva la convenzione reale dell'originale: per funzioni complesse il prodotto scalare richiede il coniugato della funzione di base e il quadrato va inteso come modulo quadrato. I riferimenti allo stato delle ricerche restano riferiti all'epoca dell'opera.

CAPITOLO VII

Curve e superfici

A. D. Aleksandrov

§1. Oggetti e metodi della teoria delle curve e delle superfici

Nei corsi scolastici la geometria considera soltanto le curve più semplici: rette, spezzate, circonferenze e archi di circonferenza; quanto alle superfici, si limita a piani, superfici di poliedri, sfere, coni e cilindri. Nei corsi più ampi si studiano altre curve, soprattutto le sezioni coniche: ellissi, parabole e iperboli. Ma lo studio di una curva o di una superficie arbitraria è del tutto estraneo alla geometria elementare. A prima vista non è nemmeno chiaro quali proprietà generali si possano scegliere come oggetto di indagine quando si parla di curve e superfici arbitrarie. Eppure un'indagine del genere è del tutto naturale e necessaria.

In ogni attività pratica e nell'osservazione della natura incontriamo continuamente curve e superfici delle forme più diverse. La traiettoria di un pianeta nello spazio, di una nave sul mare o di un proiettile nell'aria; la traccia di uno scalpello sul metallo, di una ruota sulla strada o di una penna sul nastro di uno strumento registratore; il profilo di una camma che comanda le valvole di un motore; i contorni di un disegno artistico; la forma di una fune sospesa o di una molla elicoidale avvolta per uno scopo particolare: gli esempi non finiscono mai. Le superfici degli oggetti, i gusci sottili, le cisterne, le strutture degli aeroplani, gli involucri e i materiali in fogli offrono una varietà inesauribile di superfici. I procedimenti di lavorazione dei prodotti, le proprietà ottiche degli oggetti, la loro forma aerodinamica, la rigidità o deformabilità dei gusci sottili: questi e molti altri aspetti dipendono in larga misura dalla forma geometrica delle superfici.

Naturalmente il solco lasciato da uno scalpello sul metallo non è una curva matematica. Una cisterna, anche a pareti sottili, non è una superficie matematica. Ma in prima approssimazione, sufficiente per studiare molti problemi, gli oggetti reali possono essere rappresentati matematicamente da curve e superfici.

Introducendo il concetto di curva matematica, prescindiamo da tutte le ragioni che ci impediscono di ridurre indefinitamente lo spessore. Mediante questo concetto astratto rappresentiamo quelle proprietà, del tutto concrete, dell'oggetto che si conservano quando il suo spessore e la sua larghezza diminuiscono rispetto alla lunghezza.

Analogamente, se prescindiamo dai limiti entro i quali possiamo ridurre lo spessore di un guscio o determinare con precisione i confini effettivi di un oggetto, giungiamo al concetto di superficie matematica. Non daremo una descrizione rigorosa di questi concetti ben noti;

osserviamo soltanto che le definizioni matematiche esatte non sono semplici e appartengono alla topologia.

Infine, un'importante fonte d'interesse per le curve e le superfici è stata lo sviluppo dell'analisi matematica. Basta ricordare, per esempio, che una curva è la rappresentazione geometrica di una funzione, il concetto fondamentale dell'analisi. Del resto, i grafici sono familiari a tutti anche indipendentemente dallo studio dell'analisi.

La geometria elementare creata dagli antichi Greci non trattava curve o superfici arbitrarie. Già nella geometria analitica elementare, invece, siamo abituati a dire che «ogni curva è rappresentata da un'equazione» o che «ogni equazione nelle due variabili x e y rappresenta una curva nel piano coordinato». Analogamente le superfici vengono descritte da equazioni $z = f(x, y)$ oppure $F(x, y, z) = 0$. In generale il metodo delle coordinate, stabilendo uno stretto legame fra geometria elementare e analisi, permette di definire moltissime curve e superfici.

La geometria analitica, tuttavia, limitandosi ai metodi dell'algebra e della geometria elementare, non va oltre lo studio di determinati tipi di figure. Lo studio delle curve e delle superfici arbitrarie costituisce un nuovo ramo della matematica, chiamato geometria differenziale.

Bisogna riconoscere subito che la geometria differenziale impone alle curve e alle superfici alcune condizioni derivanti dai metodi dell'analisi. Ciò non limita però in modo sostanziale la varietà delle forme ammesse, perché nella grande maggioranza dei casi esse rappresentano gli oggetti reali con la precisione necessaria. Il nome stesso «geometria differenziale» ne indica i metodi: lo strumento fondamentale è il calcolo differenziale e l'oggetto principale sono le proprietà «differenziali» delle curve e delle superfici, cioè le loro proprietà «in un punto».[1] La direzione di una curva in un punto è così determinata dalla tangente e la rapidità con cui essa si incurva è descritta dalla curvatura, di cui daremo più avanti la definizione esatta. La geometria differenziale studia le proprietà di piccoli tratti di curve e di piccole porzioni di superfici; soltanto nei suoi sviluppi successivi passa alle proprietà «in grande», cioè relative all'intera figura.

Lo sviluppo della geometria differenziale è inseparabile da quello dell'analisi. Le operazioni fondamentali dell'analisi, derivazione e integrazione, hanno un significato geometrico diretto. Come ricordato nel capitolo II, derivare una funzione $f(x)$ corrisponde a tracciare la tangente alla curva

$$y = f(x).$$

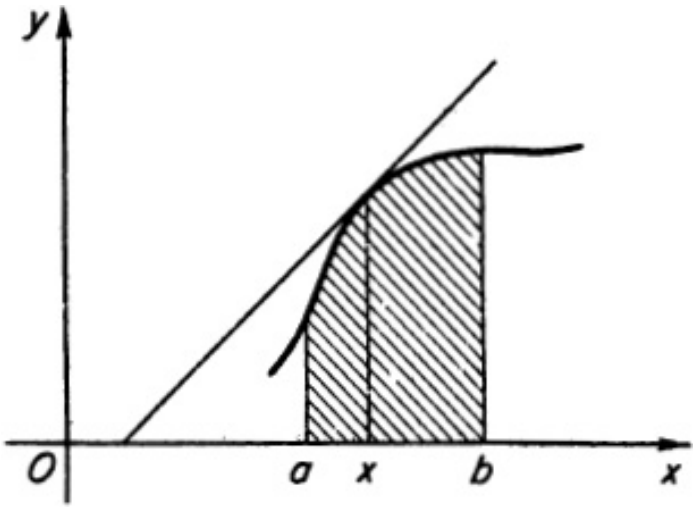


Figura 1.

Il coefficiente angolare della tangente, cioè la tangente trigonometrica dell'angolo formato con l'asse *Ox*, è precisamente la derivata $f'(x)$ nel punto corrispondente (figura 1); l'area «sotto la curva»

$$y = f(x)$$

è precisamente l'integrale $\int_a^b f(x)\ dx$ della funzione fra gli estremi corrispondenti. Come in analisi si studiano funzioni arbitrarie, così in geometria differenziale si esaminano curve e superfici arbitrarie. In analisi si comincia dall'andamento generale di una curva piana: crescita e decrescita, maggiore o minore curvatura, verso della convessità, punti di flesso e così via. Lo stretto legame fra analisi e curve è indicato dal titolo del primo manuale di analisi, del matematico francese l'Hôpital, pubblicato nel 1696: «Analisi degli infinitamente piccoli per lo studio delle linee curve».

Verso la metà del XVIII secolo, i successori immediati di Newton e Leibniz avevano sviluppato il calcolo differenziale e integrale abbastanza da aprire la strada ad applicazioni più profonde alla geometria. È soltanto da quel momento che si può parlare propriamente di una teoria delle curve e delle superfici. Per le superfici e per le curve nello spazio, i problemi analoghi sono immensamente più ricchi che per le curve piane. Col tempo essi superarono quindi il quadro di una semplice applicazione dell'analisi alla geometria e diedero origine a una teoria autonoma. Nella seconda metà del XVIII secolo molti matematici contribuirono a costruirne gli elementi: Clairaut, Eulero, Monge e altri. Fra loro Eulero va considerato il fondatore della teoria generale delle superfici. La prima trattazione complessiva delle curve e delle superfici fu l'opera di Monge «Applicazione dell'analisi alla geometria», le cui prime lezioni risalgono al 1795.[2] Dalle ricerche di questi matematici, e in particolare dall'opera di Monge, si comprendono bene le ragioni del crescente interesse per la geometria differenziale: le esigenze della meccanica, della fisica e dell'astronomia, e dunque, in ultima analisi, quelle della tecnica e dell'industria, alle quali i risultati disponibili della geometria elementare non bastavano affatto.

Anche l'opera classica di Gauss (1777-1855) sulla teoria delle superfici è legata a problemi pratici. Le sue «Ricerche generali sulle superfici curve», pubblicate nel 1827, sono fondamentali per la geometria differenziale delle superfici come disciplina autonoma. I suoi metodi e problemi generali, che discuteremo nel §4, nacquero in larga misura dalle esigenze della cartografia. Il problema cartografico consiste nel rappresentare sul piano porzioni della superficie terrestre nel modo più esatto possibile. Una rappresentazione completamente esatta è impossibile: la curvatura terrestre impone deformazioni nei rapporti fra le lunghezze. Si tratta dunque di trovare i metodi che si avvicinano maggiormente all'esattezza. La costruzione delle carte risale all'antichità più remota, ma una teoria generale è una conquista dei tempi moderni e non sarebbe stata possibile senza la teoria generale delle superfici e i metodi dell'analisi matematica. Ricordiamo che P. L. Čebyšëv (1821-1894) studiò uno dei difficili problemi matematici della cartografia, ottenendo importanti risultati sui reticoli di linee curve tracciati sulle superfici. Anche le sue indagini nacquero da problemi puramente pratici.

I problemi generali della deformazione di una superficie in modo da applicarla su un'altra costituiscono tuttora uno dei rami principali della geometria. Importanti risultati in questa direzione furono ottenuti nel 1838 da F. Minding (1806-1885), professore all'Università di Dorpat, oggi Tartu.

Nella seconda metà dell'Ottocento, la teoria delle curve e delle superfici era ormai consolidata nei suoi tratti essenziali, se ci riferiamo alla «geometria differenziale classica» e la distinguiamo dalle nuove direzioni discusse nel §5. Erano già state ottenute le equazioni fondamentali della teoria delle curve, le cosiddette formule di Frenet. Nel 1853 K. M. Peterson (1828-1881), allievo di Minding all'Università di Tartu, scoprì e studiò nella propria dissertazione le equazioni fondamentali della teoria delle superfici, ritrovate e pubblicate circa quindici anni dopo dal matematico italiano Codazzi, al cui nome sono solitamente associate. Dopo la laurea a Tartu, Peterson visse e lavorò a Mosca come insegnante in un ginnasio. Pur non avendo mai ricoperto una carica accademica adeguata ai suoi notevoli risultati scientifici, fu tra i fondatori della Società matematica di Mosca e della rivista «Matematičeskij Sbornik», pubblicata a Mosca a partire dal 1866. Con Peterson ha inizio la scuola moscovita di geometria differenziale.

I risultati della geometria differenziale «classica» furono raccolti dal geometra francese Darboux nelle sue «Lezioni sulla teoria generale delle superfici», in quattro volumi, pubblicati dal 1887 al 1896. Nel Novecento lo studio della geometria differenziale classica prosegue, ma l'interesse per curve e superfici si sposta in larga misura verso nuove direzioni, nelle quali la classe delle figure considerate si amplia ulteriormente.

§2. La teoria delle curve

Diversi modi di definire le curve in geometria differenziale.

Dall'analisi e dalla geometria analitica siamo abituati a definire le curve mediante equazioni. In un sistema cartesiano ortogonale del piano, una curva può essere data dall'equazione

$$y = f(x),$$

oppure dall'equazione più generale

$$F(x, y) = 0.$$

Questo modo di definizione si adatta però soltanto alle curve piane. Occorre anche un modo di scrivere le equazioni delle curve spaziali, che non giacciono in alcun piano. Un esempio è l'elica (figura 2).

Per la geometria differenziale, e per molti altri problemi, è particolarmente comodo rappresentare una curva come la traccia del moto continuo di un punto. La curva può naturalmente essersi formata in tutt'altro modo, ma possiamo sempre immaginarla come il cammino di un punto che la percorre.

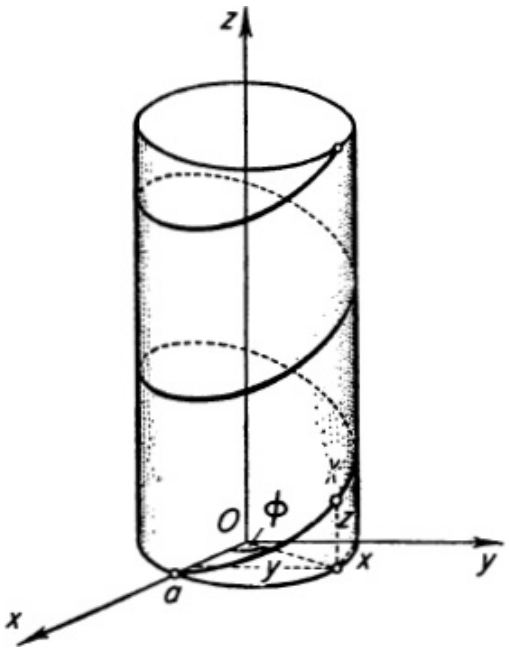


Figura 2.

Fissiamo un sistema di coordinate cartesiane nello spazio. Se un punto mobile X descrive una curva dal tempo $t = a$ al tempo $t = b$, le sue coordinate sono funzioni del tempo: $x(t)$, $y(t)$, $z(t)$. Il volo di un aeroplano o di un proiettile ne sono esempi. Viceversa, assegnate queste tre funzioni, possiamo considerarle come le coordinate di un punto mobile che descrive una curva. Le curve nello spazio possono quindi essere date da tre equazioni della forma

$$x = x(t), \quad y = y(t), \quad z = z(t).$$

Analogamente, una curva piana è definita da due equazioni

$$x = x(t), \quad y = y(t).$$

È il modo più generale di definire le curve considerate qui.

Come esempio consideriamo l'elica. Essa è prodotta dal moto di un punto che ruota uniformemente intorno a una retta, l'asse dell'elica, e nello stesso tempo si muove uniformemente in direzione parallela a essa. Prendiamo l'asse dell'elica come asse *Oz* e supponiamo che al tempo t = 0 il punto si trovi sull'asse *Ox*. Vogliamo determinare la dipendenza delle coordinate dal tempo. Se la velocità del moto parallelo a *Oz* è *c*, la distanza percorsa in questa direzione al tempo *t* è evidentemente

$$z = ct.$$

Inoltre, se ϕ è l'angolo di rotazione intorno a *Oz* e *a* la distanza del punto dall'asse, dalla figura 2 si vede che

$$x = a\cos\phi, \quad y = a\sin\phi.$$

Poiché la rotazione è uniforme, l'angolo è proporzionale al tempo: $\phi = \omega t$, dove ω è la velocità angolare. Otteniamo così

$$x = a\cos\omega t, \quad y = a\sin\omega t, \quad z = ct.$$

Queste sono le equazioni dell'elica che il punto mobile descrive al variare di *t*.

Naturalmente la variabile *t*, o parametro, come viene abitualmente chiamata, non deve necessariamente rappresentare il tempo. Possiamo inoltre sostituirla con un'altra: per esempio introdurre *u* mediante $t = u^3$ o, in generale, $t = f(u)$.[3] In geometria la scelta più naturale è la lunghezza *s* dell'arco della curva misurata da un punto fisso *A*. A ogni valore di *s* corrisponde un arco *AX*. La posizione di *X* è dunque determinata dal valore di *s* e le sue coordinate sono funzioni della lunghezza d'arco:

$$x = x(s), \quad y = y(s), \quad z = z(s).$$

Tutti questi modi di definire le curve, e altri ancora possibili,[4] consentono il calcolo numerico. Una volta definite mediante equazioni, le proprietà delle curve possono essere studiate con l'analisi matematica.

Nella geometria differenziale delle curve piane vi sono tre concetti fondamentali: lunghezza, tangente e curvatura. Per le curve spaziali si aggiungono il piano osculatore e la torsione. Ne spiegheremo ora il significato e l'importanza.

Lunghezza.

Tutti abbiamo un'idea intuitiva di ciò che significa lunghezza; occorre però trasformarla in una definizione esatta della lunghezza di una curva matematica, con un preciso contenuto numerico, che permetta di calcolarla con qualsiasi grado di accuratezza e quindi di ragionare rigorosamente sulle lunghezze. Lo stesso vale per tutti i concetti matematici. Il passaggio dalle idee informali alle misure e alle definizioni esatte segna il passaggio dalla conoscenza prescientifica degli oggetti a una teoria scientifica. L'esigenza di definire precisamente la lunghezza nacque, in ultima analisi, dalle necessità della tecnica e delle scienze naturali, il cui sviluppo richiedeva lo studio delle proprietà delle lunghezze, delle aree e di altre grandezze geometriche.

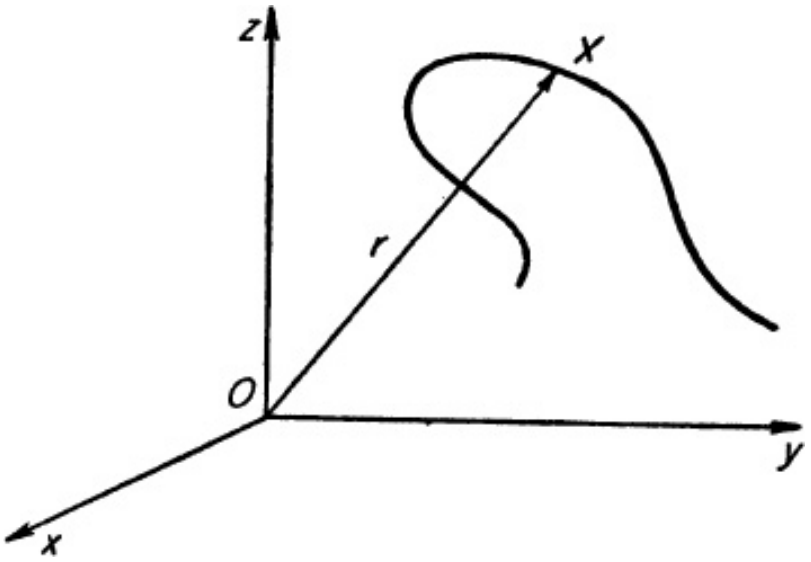


Figura 3.

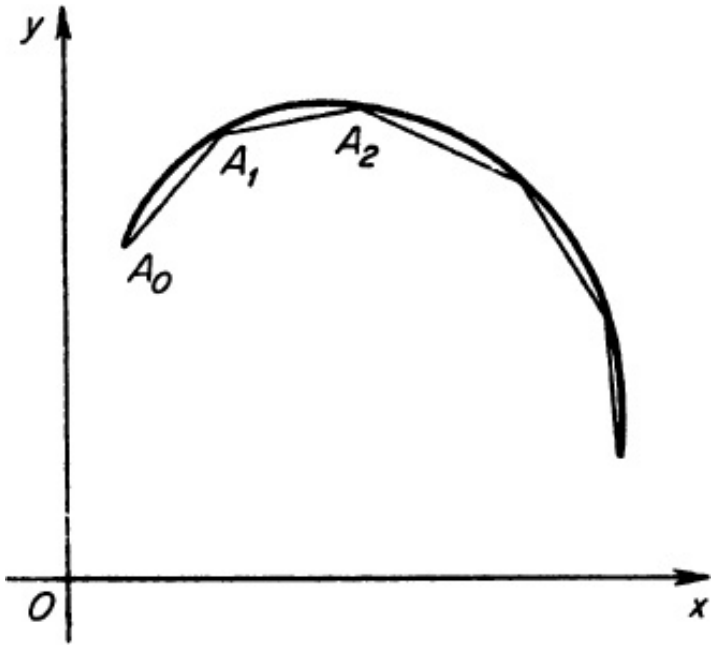


Figura 4.

Una definizione semplice e molto utile è la seguente: la lunghezza di una curva è il limite delle lunghezze delle spezzate inscritte, quando i loro vertici diventano sempre più fitti lungo la curva.

Questa definizione nasce naturalmente dai procedimenti di misura quotidiani. Scegliamo sulla curva una successione di punti $A_0, A_1, A_2, \ldots$ (figura 4) e misuriamo le distanze fra punti consecutivi. La loro somma, cioè la lunghezza della spezzata, esprime approssimativamente la lunghezza della curva. Per determinarla più esattamente è naturale avvicinare i punti, affinché la spezzata ne segua meglio le pieghe. Infine la lunghezza esatta è definita come il limite di queste approssimazioni quando i punti vengono scelti arbitrariamente vicini.[5] La definizione è dunque una generalizzazione, mediante suddivisioni sempre più fini, di un procedimento di misura completamente pratico.

Da questa definizione si ricava facilmente una formula per calcolare le lunghezze quando la curva è data analiticamente. Le formule matematiche, però, non servono soltanto al calcolo: esprimono sinteticamente teoremi che stabiliscono legami fra enti matematici diversi. Il significato teorico di tali legami può superare di molto l'utilità computazionale della formula. Per esempio, l'importanza del teorema di Pitagora, espresso dalla formula

$$c^2 = a^2 + b^2,$$

non si limita al calcolo del quadrato dell'ipotenusa c, ma sta soprattutto nel rapporto che esso stabilisce fra i lati di un triangolo rettangolo.

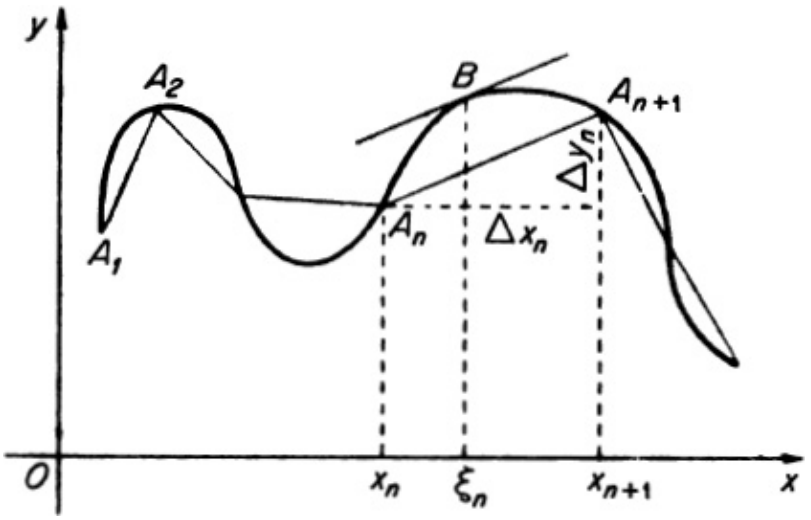


Figura 5.

Ricaviamo ora una formula per la lunghezza di una curva piana data in coordinate cartesiane da $y = f(x)$, supponendo che f abbia derivata prima continua.

Inscriviamo nella curva una spezzata (figura 5). Siano A_n, A_{n+1} due vertici consecutivi, con coordinate x_n, y_n e x_{n+1}, y_{n+1}. Il segmento che li congiunge è l'ipotenusa di un triangolo rettangolo i cui cateti sono

$$\Delta x_n = |\, x_{n+1} - x_n \,|, \quad \Delta y_n = |\, y_{n+1} - y_n \,|.$$

Per il teorema di Pitagora,

$$\overline{A_n A_{n+1}} = \sqrt{(\Delta x_n)^2 + (\Delta y_n)^2} = \sqrt{1 + \left(\frac{\Delta y_n}{\Delta x_n}\right)^2}\, \Delta x_n \,.$$

Traslando parallelamente a se stessa la retta che passa per A_n e A_{n+1}, si vede che essa assume una posizione tangente alla curva in un punto B. Sull'arco considerato esiste dunque almeno un punto in cui la tangente è parallela alla corda. Questa conclusione intuitiva può essere dimostrata rigorosamente con il teorema del valor medio.

Possiamo quindi sostituire al rapporto $\Delta y_n/\Delta x_n$ il coefficiente angolare della tangente in B, cioè $y'(\xi_n)$, dove ξ_n è l'ascissa di B. La lunghezza di un lato della spezzata diventa

$$\overline{A_n A_{n+1}} = \sqrt{1 + y'^2(\xi_n)}\, \Delta x_n \,.$$

La lunghezza totale della spezzata è la somma delle lunghezze dei lati. Indicando la somma con il simbolo Σ, abbiamo

$$S_n = \sum \sqrt{1 + y'^2(\xi_n)}\, \Delta x_n \,.$$

Per ottenere la lunghezza della curva passiamo al limite quando il massimo dei valori Δx_n tende a zero:

$$s = \lim_{\Delta x \to 0} \sum \sqrt{1 + y'^2(\xi_n)}\, \Delta x_n \,.$$

Ma questo limite è precisamente l'integrale definito nel capitolo II, ossia l'integrale della funzione $\sqrt{1 + y'^2}$. La lunghezza di una curva piana è quindi espressa dalla formula

$$s = \int_a^b \sqrt{1 + y'^2}\, dx, \qquad (1)$$

dove gli estremi d'integrazione a e b sono i valori di x alle estremità dell'arco.

La formula corrispondente per la lunghezza di una curva spaziale, un po' diversa, si ricava sostanzialmente nello stesso modo.

Il calcolo effettivo mediante queste formule non è, naturalmente, sempre semplice. Per esempio, ricavare la lunghezza di una circonferenza dalla (1) è piuttosto laborioso. Come abbiamo detto, però, l'interesse delle formule non si esaurisce nel calcolo: la (1) è importante anche per studiare le proprietà generali della lunghezza, le sue relazioni con altri concetti e così via. Avremo occasione di usarla nel capitolo VIII.

Tangente.

La tangente a una curva piana è già stata considerata nel capitolo II. Per una curva spaziale il significato è del tutto analogo. Per definire la tangente in un punto *A*, scegliamo sulla curva un punto *X* distinto da *A* e consideriamo la secante *AX*. Facciamo poi avvicinare *X* ad *A* lungo la curva. Se la secante tende a una posizione limite, la retta in tale posizione si chiama tangente nel punto *A*.[6]

Distinguendo un punto iniziale e uno finale della curva, stabiliamo un ordine di percorrenza e possiamo dire quale dei punti *A* e *X* venga prima. Per esempio, per un treno che va da Mosca a Vladivostok, Omsk precede evidentemente Irkutsk. Possiamo così orientare la secante dal primo punto al secondo. Il limite delle «secanti orientate» fornisce una «tangente orientata». Nella figura 6 la freccia indica il verso di passaggio per *A*. Quando un punto si muove lungo una curva, la velocità è diretta in ogni istante lungo la tangente.

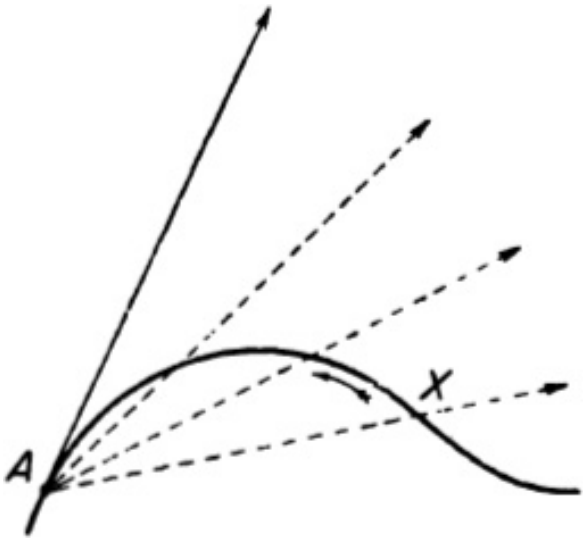


Figura 6.

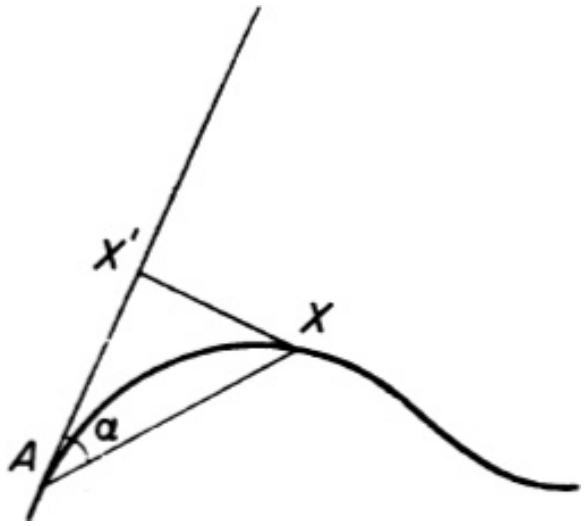


Figura 7.

La tangente possiede un'importante proprietà geometrica: vicino al punto di tangenza la curva si discosta da essa meno, in un senso preciso, che da ogni altra retta. La distanza dei punti della curva dalla tangente è molto piccola rispetto alla loro distanza dal punto di tangenza. Più esattamente, il rapporto XX'/AX (figura 7) tende a zero quando X si avvicina ad A.[7] Un piccolo tratto di curva può quindi essere sostituito con un segmento della tangente commettendo un errore piccolo rispetto alla lunghezza del tratto. Questo procedimento semplifica spesso le dimostrazioni, perché al limite fornisce risultati esatti.

È interessante osservare che, per una curva che non è una retta e quindi non ha una direzione nel senso elementare, abbiamo potuto definire una direzione in ogni punto associandole una retta. Il concetto di direzione è stato così ampliato e ha acquistato un significato nuovo. Esso riflette la natura effettiva del moto curvilineo: in ogni istante il punto si muove in una direzione determinata, che varia continuamente.

Curvatura.

Non occorre essere matematici per giudicare a occhio se un sentiero, una sottile asta o una linea disegnata siano più o meno incurvati. Ma anche per i problemi meccanici più semplici uno sguardo non basta: serve una descrizione quantitativa esatta. La otteniamo dando forma precisa all'idea intuitiva della curvatura come rapidità di variazione della direzione della curva.

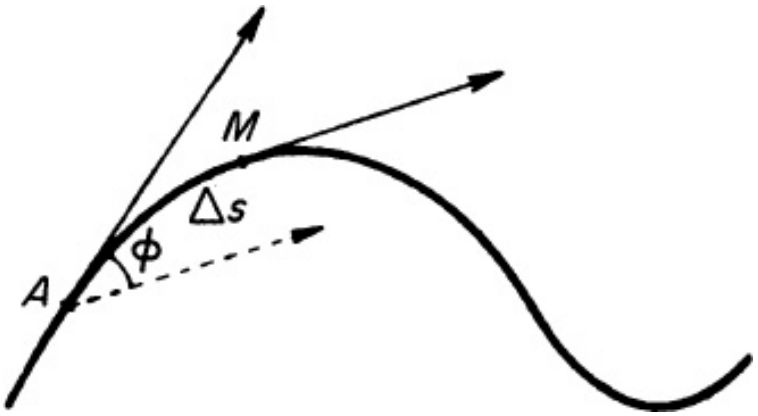


Figura 8.

Sia A un punto della curva e M un punto vicino (figura 8). L'angolo fra le tangenti nei due punti esprime quanto è cambiata la direzione nel tratto da A a M. Indichiamo tale angolo con ϕ. La variazione media di direzione per unità di lunghezza lungo l'arco AM, di lunghezza Δs, è evidentemente $\phi/\Delta s$. È quindi naturale definire la curvatura in A, cioè la rapidità di variazione della direzione nel punto stesso, come il limite di questo rapporto per $M \to A$, ossia per $\Delta s \to 0$. La curvatura è così definita dalla formula

$$k = \lim_{\Delta s \to 0} \frac{\phi}{\Delta s}.$$

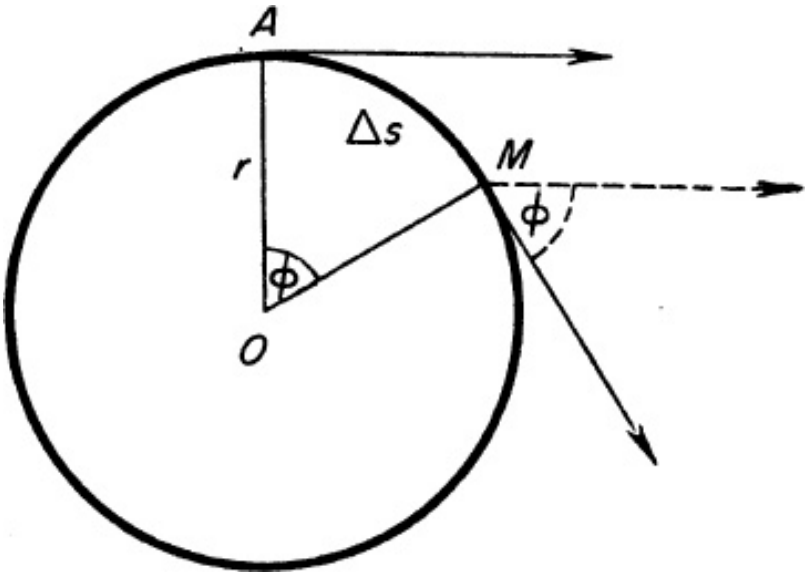


Figura 9.

Consideriamo, come esempio, la curvatura di una circonferenza (figura 9). L'angolo ϕ fra i raggi *OA* e *OM* è uguale a quello fra le tangenti in *A* e *M*, poiché le tangenti sono perpendicolari ai raggi. L'arco *AM* che sottende l'angolo ϕ ha lunghezza $\Delta s = \phi r$, perciò

$$\frac{\phi}{\Delta s} = \frac{1}{r}.$$

Il rapporto $\phi/\Delta s$ è dunque costante e la curvatura della circonferenza, suo valore limite, è in ogni punto uguale al reciproco del raggio.[8]

Ricaviamo la formula della curvatura di una curva piana data da $y = f(x)$. Come origine della lunghezza d'arco scegliamo un punto fisso *N* (figura 10). L'angolo ϕ fra le tangenti in *A* e *M* è, in valore assoluto, la differenza dei loro angoli d'inclinazione:

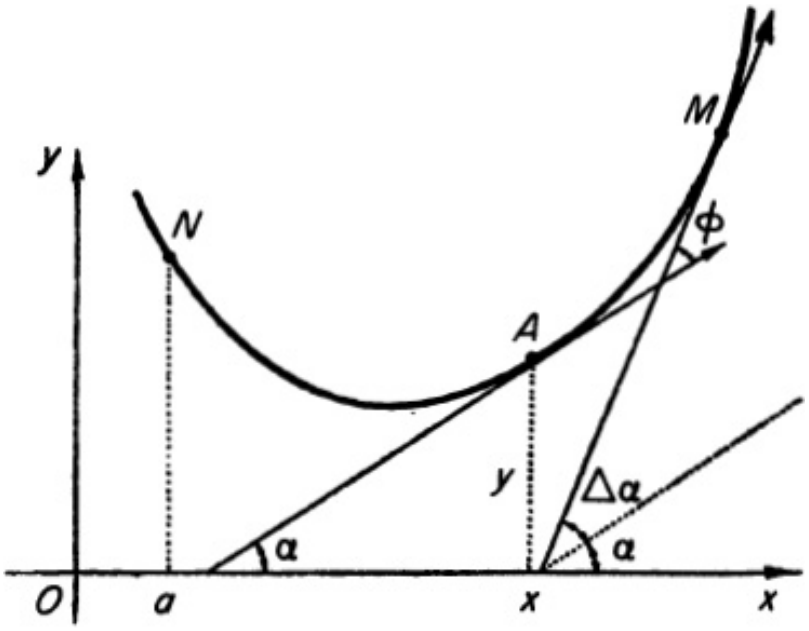


Figura 10.

$$\phi = |\Delta\alpha|.$$

Poiché l'angolo α può diminuire, prendiamo il valore assoluto $|\Delta\alpha|$. Ci interessa il valore

$$k = \lim_{\Delta s \to 0} \frac{\phi}{\Delta s} = \lim_{\Delta s \to 0} \frac{|\Delta\alpha|}{\Delta s} = \lim_{\Delta x \to 0} \frac{\frac{|\Delta\alpha|}{\Delta x}}{\frac{\Delta s}{\Delta x}} = \frac{|\alpha'|}{s'}.$$

La lunghezza dell'arco NA è espressa dall'integrale

$$s = \int_a^x \sqrt{1 + y'^2}\, dx,$$

quindi

$$s' = \sqrt{1 + y'^2}.$$

Resta da trovare α'. Sappiamo che $\tan \alpha = y'$ e dunque, localmente, $\alpha = \arctan y'$, a meno di una costante multipla di π. Derivando rispetto a x, otteniamo

$$\alpha' = \frac{1}{1 + y'^2}\, y''.$$

Infine,

$$k = \frac{|\alpha'|}{s'} = \frac{|y''|}{(1 + y'^2)^{3/2}}.$$

Le formule corrispondenti ad altri modi di rappresentare curve piane e spaziali si trovano nei consueti corsi di analisi o di geometria differenziale.

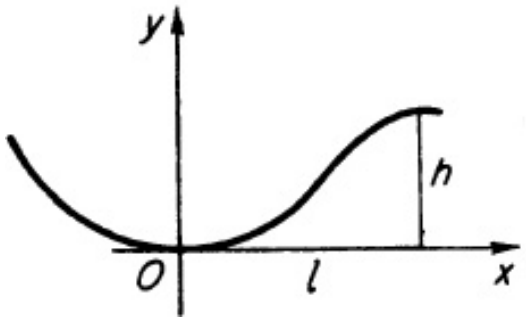


Figura 11.

Questa formula permette un'altra interpretazione geometrica della curvatura, utile in molti problemi. La curvatura di una curva in un punto è infatti data da

$$k = \lim_{l \to 0} \frac{2h}{l^2},$$

dove h è la distanza di un secondo punto della curva dalla tangente nel punto dato e l è la lunghezza del segmento di tangente compreso fra il punto di tangenza e la proiezione del secondo punto sulla tangente (figura 11).

Per dimostrarlo scegliamo un sistema cartesiano ortogonale con origine nel punto dato e asse Ox tangente alla curva (figura 11). Per semplicità supponiamo che la curva sia piana. Nell'origine $y' = 0$ e $k = |y''|$. Sviluppando $y = f(x)$ mediante la formula di Taylor, otteniamo $y = \frac{1}{2} y'' x^2 + \epsilon x^2$, avendo tenuto conto che $y' = 0$. Qui $\epsilon \to 0$ per $x \to 0$. Ne segue $k = |y''| = \lim_{x \to 0} 2|y|/x^2$; essendo $|y| = h$ e $x^2 = l^2$, abbiamo

$$k = \lim_{l \to 0} \frac{2h}{l^2}.$$

Questa formula mostra che la curvatura descrive la rapidità con cui la curva si allontana dalla tangente.

Passiamo ad alcune applicazioni molto importanti della curvatura alla meccanica.

Consideriamo anzitutto il seguente problema. Una fune flessibile sia tesa su un sostegno (figura 12) in modo da rimanere in un piano. Vogliamo determinare in ogni punto la pressione della fune sul sostegno o, più precisamente, il limite

$$p = \lim_{\Delta s \to 0} \frac{P}{\Delta s}, \qquad (2)$$

dove P è l'intensità della forza vettoriale P esercitata sul sostegno lungo un tratto di lunghezza Δs contenente il punto dato. Per semplicità supponiamo che il modulo T della tensione sia uguale in tutti i punti della fune.

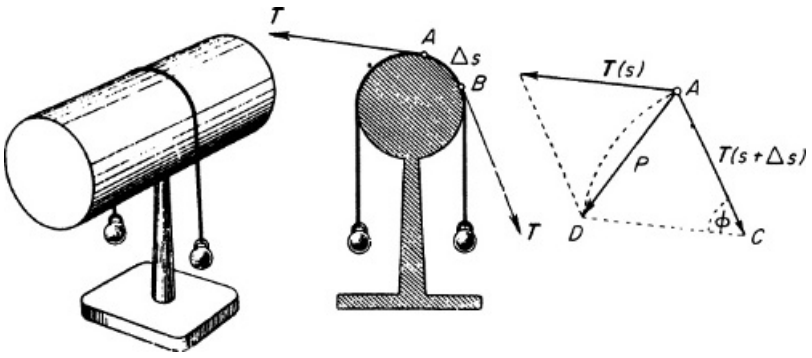


Figura 12.

Consideriamo il punto A e un tratto AB della fune.[9] Sul tratto di lunghezza Δs, oltre alla reazione del sostegno, agiscono soltanto due forze esterne: le tensioni agli estremi, uguali in modulo e dirette lungo le rispettive tangenti. La forza P esercitata dalla fune sul sostegno è quindi uguale alla somma vettoriale delle tensioni agli estremi. Come mostra la figura 12, il

vettore P è rappresentato dalla base AD del triangolo isoscele CAD. I lati uguali hanno lunghezza T e l'angolo al vertice C è uguale alla variazione di direzione della tangente da A a B.

Al diminuire di Δs, l'angolo ϕ diminuisce e l'angolo fra P e la tangente in A tende a un angolo retto. La pressione è dunque perpendicolare alla tangente.

Per trovarne l'intensità usiamo il fatto che un piccolo arco di circonferenza ha lunghezza approssimativamente uguale alla corda che lo sottende.

Sostituiamo quindi alla lunghezza della corda AD, ossia a P, la lunghezza $T\phi$ dell'arco corrispondente. Dalla (2) otteniamo

$$p = \lim_{\Delta s \to 0} \frac{P}{\Delta s} = \lim_{\Delta s \to 0} \frac{T\phi}{\Delta s} = T \lim_{\Delta s \to 0} \frac{\phi}{\Delta s} = Tk.$$

La pressione per unità di lunghezza è dunque, in ogni punto, il prodotto della curvatura per la tensione della fune ed è perpendicolare alla tangente.

Consideriamo un secondo problema. Un punto materiale, cioè un corpo molto piccolo, si muova lungo una curva piana con velocità di modulo costante v. Qual è la sua accelerazione in un punto A? Per definizione l'accelerazione è il limite del rapporto fra la variazione della velocità durante il tempo Δt e Δt. La velocità comprende non soltanto un modulo, ma anche una direzione: dobbiamo quindi considerare la variazione del vettore velocità. Il problema matematico di trovare il modulo dell'accelerazione consiste nel calcolare il limite

$$w = \lim_{\Delta t \to 0} \frac{|\, \boldsymbol{v}(t + \Delta t) - \boldsymbol{v}(t)\,|}{\Delta t},$$

dove $v(t)$ è il vettore velocità nel punto A e $|v(t + \Delta t) - v(t)|$ è il modulo della differenza fra i vettori velocità. Il limite può essere scritto anche come

$$\lim_{\Delta s \to 0} \frac{|\, -\boldsymbol{v}(t) + \boldsymbol{v}(t + \Delta t)\,|}{\Delta s} \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t},$$

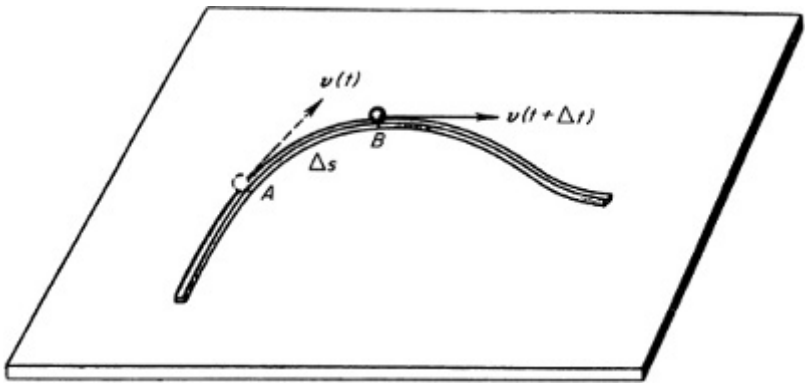


Figura 13.

dove Δs è la lunghezza dell'arco AB percorso nel tempo Δt. Guardando la figura 13 e ricordando che la velocità è sempre tangente alla curva e ha modulo costante, vediamo che

trovare la somma $-v(t) + v(t + \Delta t)$ equivale geometricamente a trovare il vettore P del problema precedente. Possiamo quindi usare quel risultato e, sostituendo la velocità alla tensione, scrivere

$$\lim_{\Delta s \to 0} \frac{| -v(t) + v(t + \Delta t) |}{\Delta s} = vk.$$

Inoltre $\lim_{\Delta t \to 0} \Delta s/\Delta t = v$. Otteniamo così che l'accelerazione di un corpo in moto uniforme lungo una curva è uguale al prodotto della curvatura per il quadrato della velocità:

$$w = kv^2 \qquad (3)$$

ed è diretta lungo la normale alla curva, cioè lungo una retta perpendicolare alla tangente.

Il ricorso a un'analogia geometrica, che ci ha permesso di usare la soluzione del problema della pressione di una fune per risolvere quello dell'accelerazione di una particella, mostra ancora una volta quanto sia utile astrarre dalle proprietà concrete particolari di un fenomeno e passare ai concetti e ai risultati matematici corrispondenti: questi ultimi possono poi essere applicati alle situazioni più diverse.

Osserviamo inoltre che la curvatura, che dal punto di vista meccanico esprime la variazione della direzione del moto, è strettamente collegata alle forze che producono tale variazione. Moltiplicando la (3) per la massa m del punto mobile, otteniamo facilmente la relazione

$$F_n = mw = v^2mk.$$

Qui F_n è il modulo della componente normale della forza agente sul punto.

Piano osculatore.

Benché una curva spaziale non giaccia in un piano, di regola a ogni suo punto A si può associare un piano P che, nelle vicinanze del punto, è più vicino alla curva di ogni altro piano. Esso si chiama piano osculatore della curva nel punto.

Il piano osculatore passa naturalmente per A e contiene la tangente T. Ma vi sono molti piani contenenti questa retta. Per scegliere quello che si discosta meno dalla curva, esaminiamo come la curva si allontana dalla tangente. Proiettiamola sul piano normale Q, perpendicolare a T in A (figura 14). La proiezione di un tratto della curva contenente A forma una nuova curva, indicata in figura con una linea punteggiata, che solitamente presenta una cuspide in A. Se essa ha una tangente N in questo punto, il piano P individuato da T e N è naturalmente quello più vicino alla curva originaria nelle vicinanze di A: è il suo piano osculatore. Si può dimostrare che, se le funzioni che definiscono una curva regolare possiedono derivate seconde e la curvatura in A è non nulla, il piano osculatore esiste necessariamente. La sua equazione si esprime in modo molto semplice mediante le derivate prime e seconde delle funzioni che definiscono la curva.

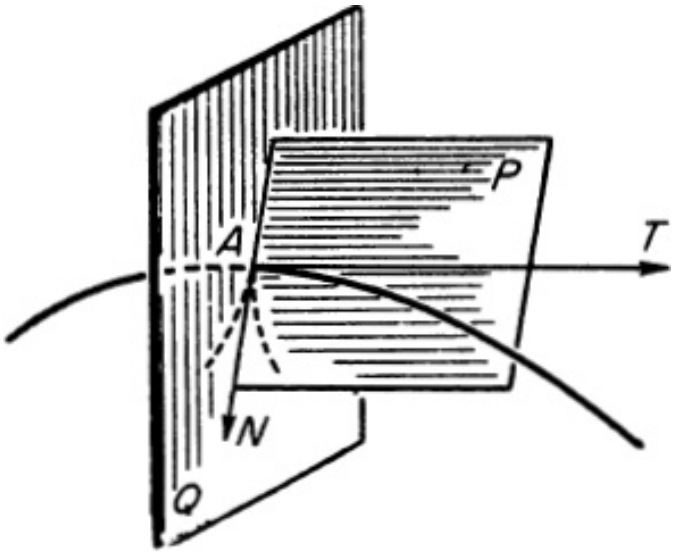


Figura 14.

Abbiamo visto che le proprietà della tangente permettono di trattare un piccolo arco di curva piana come un segmento rettilineo, con un errore piccolo rispetto alla lunghezza dell'arco. Analogamente, le proprietà del piano osculatore permettono di sostituire un piccolo arco di curva spaziale con la sua proiezione su tale piano; l'errore è allora piccolo rispetto al quadrato della lunghezza dell'arco.

Le rette dello spazio perpendicolari alla tangente nel punto dato formano il piano normale della curva. Fra esse ve n'è una, *N*, che giace nel piano osculatore: si chiama normale principale. Di solito la si orienta verso la concavità della proiezione della curva sul piano osculatore. La normale principale svolge per una curva spaziale il ruolo della normale ordinaria, unica, di una curva piana. In particolare, per una fune sottile tesa con tensione *T* su un sostegno lungo una curva spaziale, la pressione per unità di lunghezza ha modulo *Tk* ed è diretta lungo la normale principale. Se un punto materiale percorre una curva spaziale con velocità di modulo costante *v*, la sua accelerazione è kv^2 ed è diretta lungo la normale principale.

Torsione.

Passando da un punto all'altro della curva, la posizione del piano osculatore cambia in generale. Come la rapidità di variazione della tangente caratterizza la curvatura, così quella del piano osculatore caratterizza una nuova grandezza: la torsione. Anche qui la rapidità è riferita alla lunghezza d'arco. Se ψ è l'angolo orientato fra i piani osculatori in un punto fisso *A* e in un punto vicino *X*, e Δs è la lunghezza dell'arco *AX*, la torsione τ in *A* è definita dal limite[10]

$$\tau = \lim_{\Delta s \to 0} \frac{\psi}{\Delta s}.$$

Il segno della torsione dipende dal verso in cui il piano osculatore ruota mentre si procede lungo la curva, secondo una convenzione di orientazione fissata.

Possiamo immaginare il piano osculatore come una stecca di un ventaglio sulla quale siano disegnate la tangente e la normale principale. In ogni istante la tangente ruota verso la normale con una rapidità determinata dalla curvatura, mentre il piano osculatore ruota intorno alla tangente con rapidità e verso determinati dalla torsione.

I risultati più semplici della teoria delle equazioni differenziali permettono di dimostrare un teorema fondamentale che, in termini approssimativi, afferma che due curve con la stessa curvatura e la stessa torsione sono uguali. Precisiamo il significato. Procedendo lungo una curva fino a distanze d'arco s diverse dal punto iniziale, incontriamo valori diversi di k e τ, dipendenti da s. Otteniamo quindi due funzioni ben determinate della lunghezza d'arco, $k(s)$ e $\tau(s)$.

Il teorema afferma che, nelle consuete ipotesi di regolarità e con curvatura non nulla, due curve aventi la stessa curvatura e la stessa torsione come funzioni di s sono congruenti: una può essere portata a coincidere con l'altra mediante un movimento rigido. Curvatura e torsione in funzione della lunghezza d'arco determinano dunque completamente la curva, salvo la sua posizione nello spazio. Esse ne descrivono le proprietà stabilendo le relazioni fra lunghezza, curvatura e torsione. Questi tre concetti costituiscono così una sorta di fondamento ultimo per lo studio delle curve. Servono anche a esprimere i concetti più semplici della teoria delle superfici, alla quale ora passiamo.

Naturalmente le osservazioni precedenti non esauriscono la teoria delle curve. Vi sono molti altri concetti e problemi: tipi speciali di curve, famiglie di curve, loro posizione sulle superfici, forma complessiva di una curva e così via. Queste questioni e i metodi per affrontarle sono collegati a quasi tutti i rami della matematica. L'insieme dei problemi risolvibili mediante la teoria delle curve è estremamente ricco e vario.

§3. Concetti fondamentali della teoria delle superfici

I principali modi di definire una superficie.

Per studiare le superfici mediante l'analisi dobbiamo naturalmente definirle analiticamente. Il modo più semplice è un'equazione

$$z = f(x, y),$$

in cui x, y e z sono le coordinate cartesiane di un punto della superficie. La funzione $f(x, y)$ non deve necessariamente essere definita per ogni coppia x, y: il suo dominio può avere forme diverse. La superficie della figura 15, per esempio, è data da una funzione definita in una corona circolare. Esempi di superfici descritte da $z = f(x, y)$ sono già familiari dalla geometria analitica: $z = Ax + By + C$ rappresenta un piano e $z = x^2 + y^2$ un paraboloide di rivoluzione (figura 16). Per applicare il calcolo differenziale occorre che f possieda derivate prime, seconde e talvolta di ordine maggiore, con la regolarità richiesta. Una superficie così definita si dice regolare. Geometricamente ciò significa, in modo non del tutto preciso, che essa si incurva con

continuità, senza spigoli o altre singolarità. Superfici con cuspidi, spigoli o altre singolarità richiedono un nuovo tipo di studio (cfr. §5).

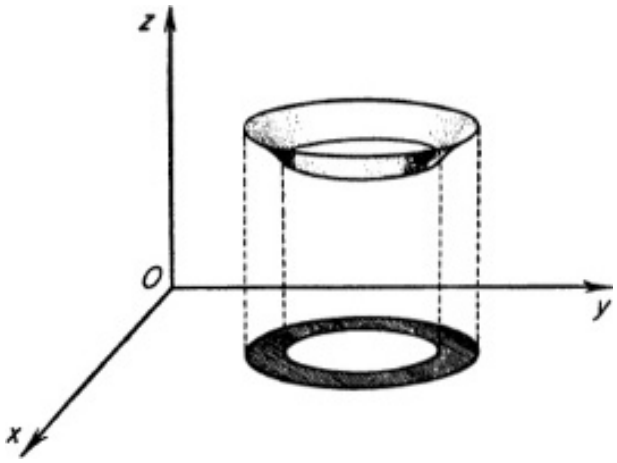


Figura 15.

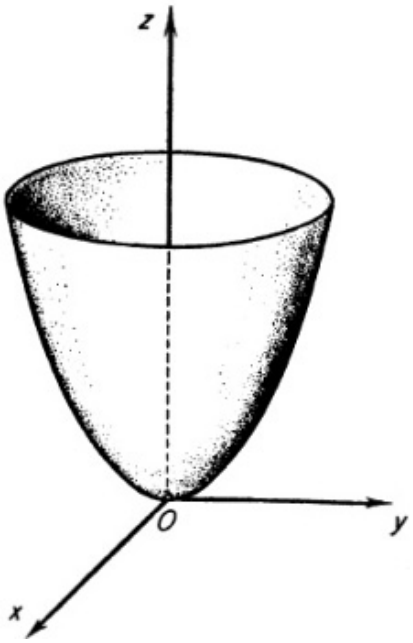


Figura 16.

Non ogni superficie, anche priva di singolarità, può però essere rappresentata interamente nella forma $z = f(x, y)$. Se a ogni coppia x, y del dominio corrisponde un unico valore di z, ogni retta parallela a Oz interseca la superficie al più in un punto (figura 17). Già superfici semplici come sfere e cilindri non sono rappresentabili globalmente in questa forma. Si ricorre allora ad altri modi, per esempio a $F(x, y, z) = 0$. Una sfera di raggio R e centro nell'origine ha così equazione

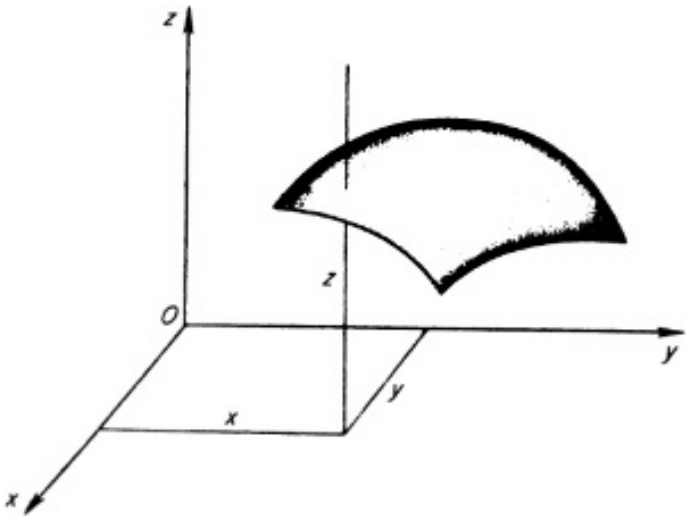


Figura 17.

$$x^2 + y^2 + z^2 = R^2.$$

L'equazione $x^2 + y^2 = r^2$ rappresenta un cilindro di raggio r.

Se lo studio riguarda soltanto piccole porzioni della superficie, come accade di solito nella geometria differenziale classica, la forma $z = f(x, y)$ è del tutto generale: scegliendo opportunamente gli assi, ogni porzione abbastanza piccola di una superficie liscia può essere rappresentata così. La prenderemo come forma fondamentale, rinviando altri modi di definizione ai §§4 e 5.

Piano tangente.

Come una curva liscia ha in ciascun punto una retta tangente che le è vicina nell'intorno del punto, così una superficie può avere in ciascun punto un piano tangente.

La definizione esatta è la seguente. Un piano P passante per un punto M di una superficie F è tangente a F in M se l'angolo α fra P e la secante MX, che congiunge M a un punto X della superficie, tende a zero quando X tende a M (figura 18). Tutte le tangenti alle curve della superficie passanti per M giacciono evidentemente nel piano tangente.

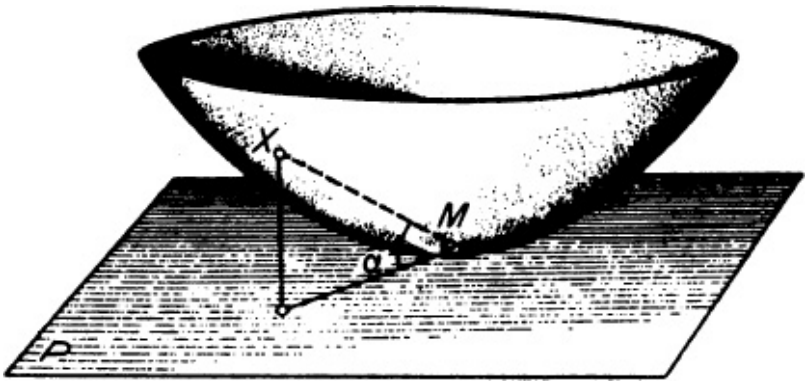


Figura 18.

Una superficie *F* si dice liscia se ha un piano tangente in ogni punto e la posizione di tale piano varia con continuità passando da un punto all'altro.

Vicino al punto di tangenza la superficie si discosta pochissimo dal piano tangente. Se *X* tende a *M* lungo la superficie, la sua distanza dal piano diventa sempre più piccola anche rispetto alla distanza da *M*, come il lettore può verificare nella figura 18. In questo senso, vicino a *M* la superficie quasi si confonde con il piano tangente. In prima approssimazione una piccola porzione, o «elemento», di superficie può essere sostituita con una porzione di quel piano. La perpendicolare al piano tangente passante per il punto di tangenza svolge il ruolo di perpendicolare alla superficie e si chiama normale.

Questa sostituzione è utile in molte situazioni. La riflessione della luce su una superficie curva, per esempio, avviene come su un piano: il raggio incidente e quello riflesso giacciono nello stesso piano della normale e formano con essa angoli uguali (figura 19), come se la riflessione avvenisse sul piano tangente. Analogamente, nella rifrazione su una superficie curva, ogni raggio segue la consueta legge di rifrazione come se l'elemento attraversato fosse piano. Questi fatti sono alla base dei calcoli di riflessione e rifrazione negli strumenti ottici. Inoltre due corpi solidi lisci a contatto hanno nel punto di contatto un piano tangente comune. In assenza di attrito, la pressione di un corpo sull'altro è diretta lungo la normale. Ciò vale anche quando i punti di contatto sono più di uno: in ciascuno la pressione segue la normale corrispondente.

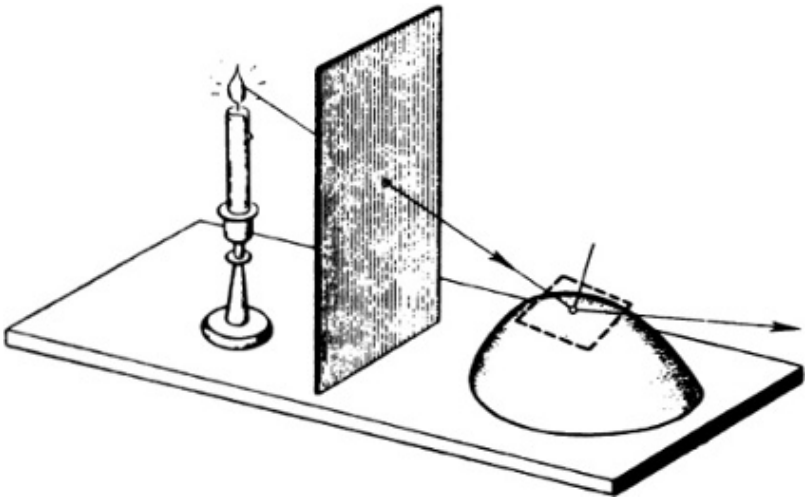


Figura 19.

La sostituzione degli elementi di superficie con porzioni dei piani tangenti può anche servire a definire l'area. Si divide la superficie in piccole parti F_1, F_2, ..., F_n e si proietta ciascuna sul piano tangente in un suo punto (figura 20). Si ottengono regioni piane P_1, P_2, ..., P_n, la somma delle cui aree approssima l'area della superficie. Quest'ultima è definita come il limite di tali somme per suddivisioni sempre più fini.[11] Se ne ricava un'espressione esatta mediante un integrale doppio.

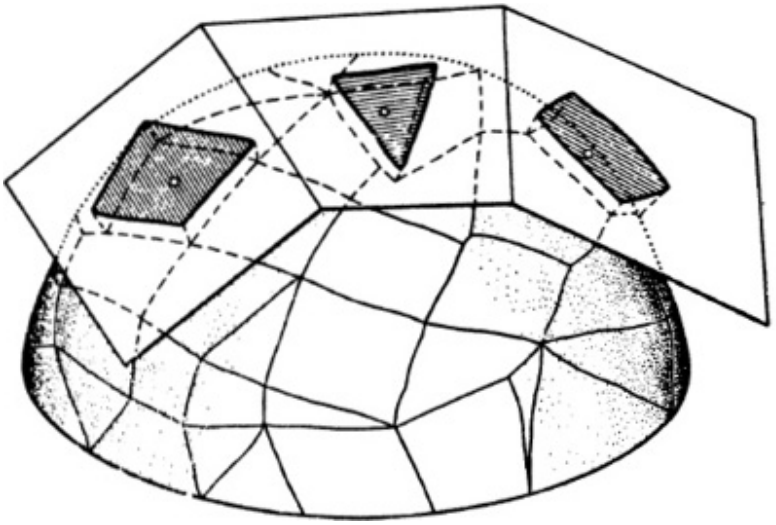


Figura 20.

Queste osservazioni mostrano chiaramente l'importanza del piano tangente. In molti problemi, però, l'approssimazione piana di un elemento di superficie non basta: occorre considerarne la curvatura.

Curvatura delle curve su una superficie.

La curvatura di una superficie in un punto è caratterizzata dalla rapidità con cui essa si allontana dal piano tangente. Ma questa rapidità può essere diversa nelle varie direzioni. La superficie della figura 21, per esempio, si allontana dal piano *P* più rapidamente nella direzione *OA* che in quella *OB*. È quindi naturale definire la curvatura della superficie mediante le curvature delle curve che vi giacciono e passano per il punto dato nelle diverse direzioni.

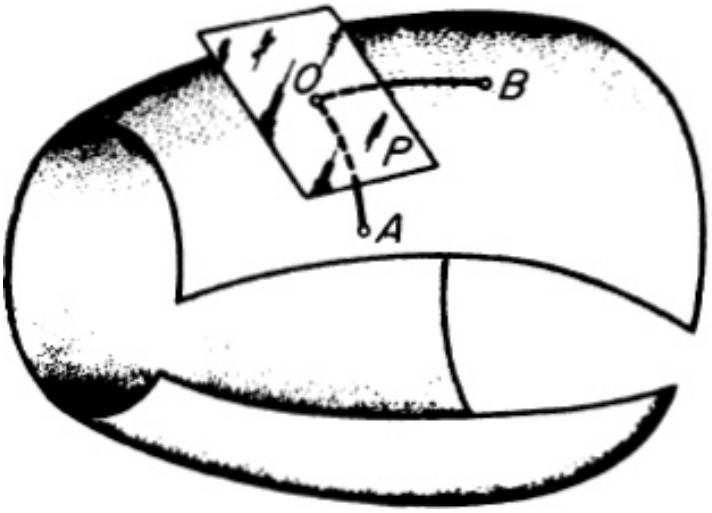


Figura 21.

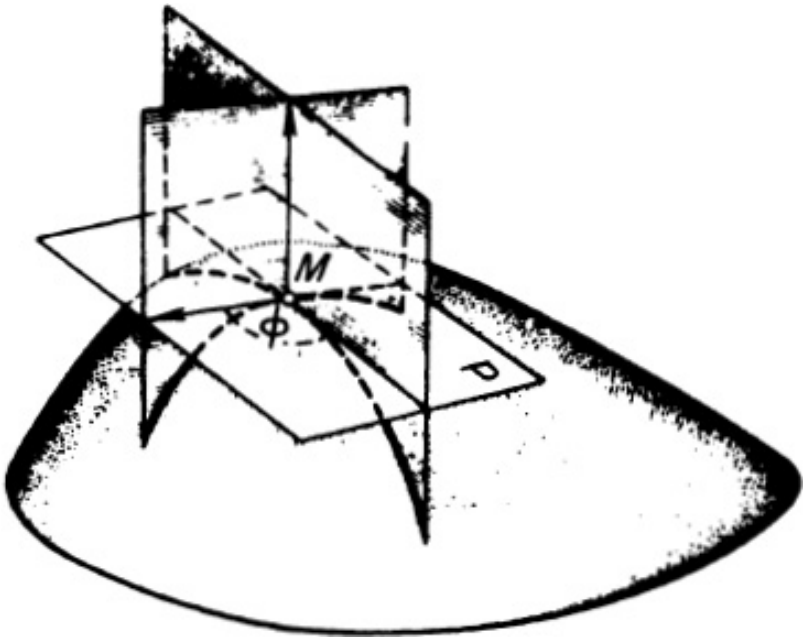


Figura 22.

Procediamo così. Costruiamo il piano tangente *P* in *M* e scegliamo un verso sulla normale (figura 22). Consideriamo poi le curve ottenute intersecando la superficie con piani passanti per la normale in *M*: si chiamano sezioni normali. Alla loro curvatura attribuiamo un segno, positivo se la concavità è rivolta nel verso della normale e negativo nel caso opposto. Sulla superficie a sella della figura 23, con la normale orientata come indica la freccia, la sezione *MA* ha curvatura positiva e la *MB* curvatura negativa.

Una sezione normale è individuata dall'angolo ϕ fra la sua direzione nel piano tangente e una semiretta iniziale fissata (figura 22). Conoscendo la curvatura $k(\phi)$ in funzione di questo angolo, abbiamo una descrizione piuttosto completa del comportamento della superficie vicino a *M*.

Una superficie può incurvarsi in molti modi e sembrerebbe quindi che la dipendenza di *k* da ϕ possa essere arbitraria. In realtà non è così. Per le superfici della geometria differenziale esiste una semplice legge, dovuta a Eulero, che collega le curvature delle sezioni normali passanti per un punto nelle diverse direzioni.

Si dimostra che in ogni punto esistono due direzioni particolari tali che:

1. Sono perpendicolari fra loro.

2. Le curvature k_1 e k_2 delle rispettive sezioni normali sono i valori minimo e massimo fra le curvature di tutte le sezioni normali, nell'uno o nell'altro ordine.[12]

3. La curvatura $k(\phi)$ della sezione la cui direzione forma l'angolo ϕ con quella di curvatura k_1 è espressa da

$$k(\phi) = k_1 \cos^2 \phi + k_2 \sin^2 \phi. \qquad (4)$$

Queste direzioni si chiamano direzioni principali e k_1, k_2 sono le curvature principali della superficie nel punto dato.

Il teorema di Eulero mostra che, nonostante la varietà delle superfici, la loro forma vicino a ogni punto appartiene, nell'approssimazione del secondo ordine, a pochi tipi ben determinati. Se k_1 e k_2 hanno lo stesso segno, anche $k(\phi)$ mantiene quel segno e la superficie ha localmente la forma della figura 22. Se hanno segno opposto, per esempio $k_1 > 0$ e $k_2 < 0$, la curvatura delle sezioni normali cambia segno: per $\phi = 0$ vale $k = k_1 > 0$, mentre per $\phi = \pi/2$ vale $k = k_2 < 0$.

Dalla formula (4) si dimostra facilmente che, al variare di ϕ da 0 a π, il segno di $k(\phi)$ cambia due volte;[13] vicino al punto la superficie ha quindi forma di sella (figura 23).

Quando una sola delle curvature principali è nulla, la curvatura mantiene lo stesso segno tranne nella direzione in cui si annulla. È quanto accade in ogni punto di un cilindro (figura 24). In generale, vicino a un punto di questo tipo la superficie ha, nell'approssimazione del secondo ordine, una forma simile a quella di un cilindro.

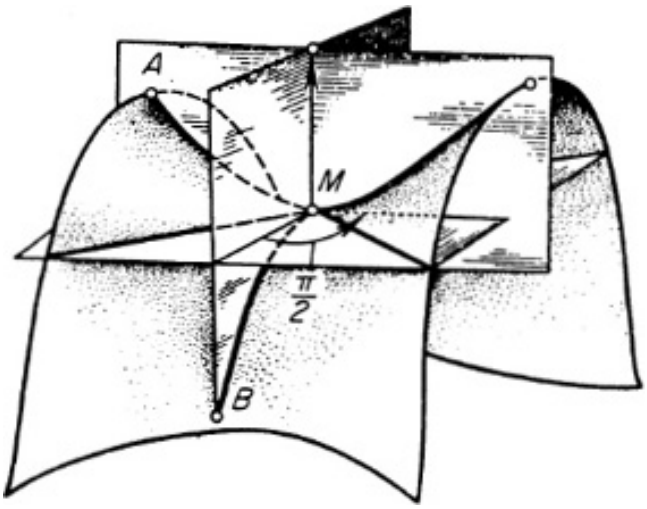


Figura 23.

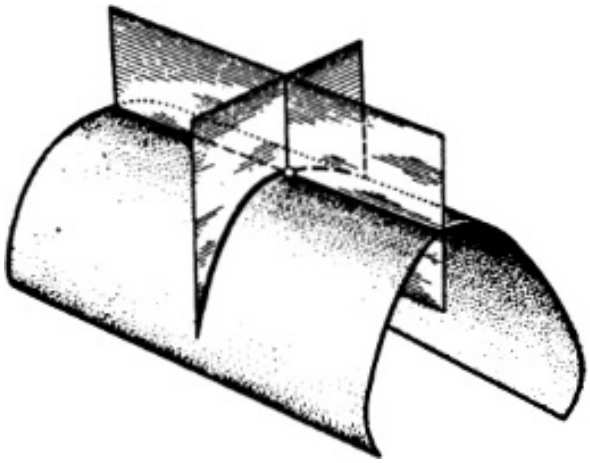


Figura 24.

Infine, se $k_1 = k_2 = 0$, tutte le sezioni normali hanno curvatura nulla. Vicino a un tale punto la superficie è particolarmente «vicina» al piano tangente. Questi punti si chiamano punti planari; un esempio è il punto *M* della figura 25. Il comportamento della superficie nelle loro vicinanze può essere assai complesso.



Figura 25.

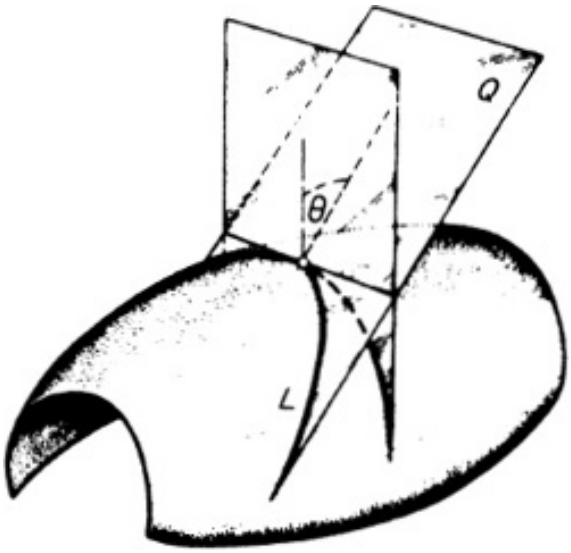


Figura 26.

Consideriamo ora una sezione prodotta da un piano *Q* che non contiene la normale e non coincide con il piano tangente (figura 26). Come mostrò Meusnier,[14] la curvatura k_L della curva risultante *L* è legata da una semplice relazione alla curvatura k_N della sezione normale nella stessa direzione, cioè quella che interseca il piano tangente lungo la stessa retta. La relazione è

$$k_L = \frac{|k_N|}{\cos\theta},$$

dove θ è l'angolo fra il piano *Q* e il piano della sezione normale corrispondente. L'esattezza della formula si visualizza molto bene su una sfera.

Si può inoltre dimostrare che una curva della superficie avente Q come piano osculatore ha, nel punto considerato, la stessa curvatura della sezione prodotta da Q, purché questo piano non sia tangente alla superficie.

Conoscendo k_1 e k_2, la curvatura di una curva sulla superficie si determina quindi dalla direzione della sua tangente e dall'inclinazione del piano osculatore rispetto al piano della sezione normale, nei casi non degeneri appena considerati. Il carattere della curvatura della superficie in un punto è dunque descritto dai due numeri k_1 e k_2. I loro valori assoluti sono le curvature di due sezioni normali perpendicolari; i segni indicano il verso della concavità rispetto all'orientazione scelta sulla normale.

Dimostriamo ora i teoremi di Eulero e di Meusnier.

1. Per il teorema di Eulero serve il lemma seguente. Se $f(x, y)$ ha derivate seconde continue vicino a un punto, si possono ruotare gli assi di un angolo α in modo che, nelle nuove coordinate, la derivata mista $f_{x'y'}$ sia nulla in quel punto.[15] Ricordiamo che dopo la rotazione le nuove variabili sono collegate alle vecchie dalle formule

$$x = x' \cos \alpha - y' \sin \alpha; \quad y = x' \sin \alpha + y' \cos \alpha$$

(cfr. capitolo III, §7). Per dimostrare il lemma osserviamo che

$$\frac{\partial x}{\partial x'} = \cos \alpha, \ \frac{\partial y}{\partial x'} = \sin \alpha, \ \frac{\partial x}{\partial y'} = -\sin \alpha, \ \frac{\partial y}{\partial y'} = \cos \alpha.$$

Calcolando $f_{x'y'}$ con la regola di derivazione delle funzioni composte, dopo qualche passaggio otteniamo

$$f_{x'y'} = f_{xy} \cos 2\alpha + \tfrac{1}{2}(f_{yy} - f_{xx}) \sin 2\alpha,$$

da cui segue facilmente che, scegliendo

$$\cot 2\alpha = \frac{1}{2} \frac{f_{xx} - f_{yy}}{f_{xy}}$$

si ha

$$f_{x'y'} = 0.$$

Consideriamo ora la superficie F, di equazione $z = f(x, y)$, con origine nel punto M in esame e assi Ox, Oy scelti nel piano tangente in modo che $f_{xy}(0, 0) = 0$. Nel piano tangente P prendiamo una retta che formi l'angolo ϕ con Ox e consideriamo la sezione normale L in quella direzione (figura 27). Dalla formula ricavata nel §2, tenendo conto del segno, la sua curvatura in M è

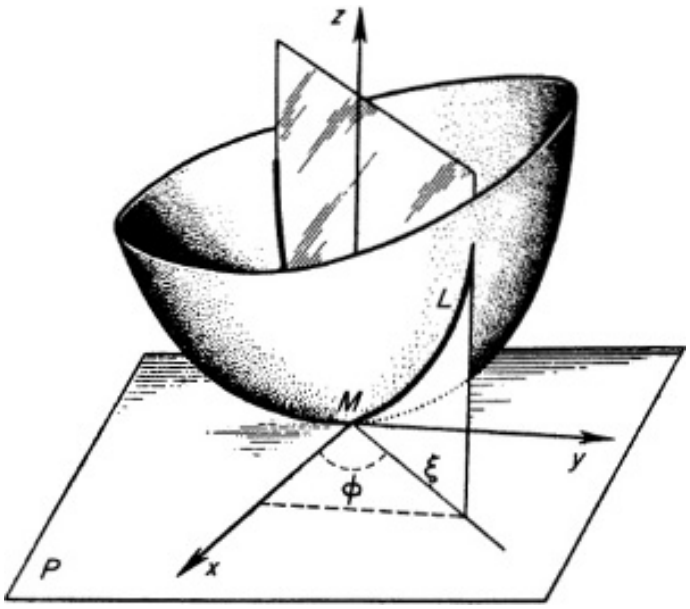


Figura 27.

$$k_L = \lim_{\xi \to 0} \frac{2f(x, y)}{\xi^2}.$$

Qui $f(x, y)$ è la distanza con segno di un punto di L dalla retta scelta. Sviluppando f mediante Taylor (capitolo II, §9) e osservando che $f_x(0, 0) = f_y(0, 0) = 0$, poiché gli assi giacciono nel piano tangente, otteniamo

$$f(x, y) = \tfrac{1}{2}(f_{xx}x^2 + f_{yy}y^2) + \epsilon(x^2 + y^2),$$

dove $\epsilon \to 0$ per $x, y \to 0$. Per un punto di L abbiamo $x = \xi \cos \phi$, $y = \xi \sin \phi$, $\xi^2 = x^2 + y^2$ (figura 27), quindi

$$k_L = \lim_{\xi \to 0} \frac{f_{xx}\xi^2 \cos^2 \phi + f_{yy}\xi^2 \sin^2 \phi + 2\epsilon\xi^2}{\xi^2} = f_{xx} \cos^2 \phi + f_{yy} \sin^2 \phi.$$

Ponendo $\phi = 0$ e $\phi = \pi/2$, troviamo che f_{xx} e f_{yy} sono le curvature k_1 e k_2 delle sezioni normali nelle direzioni degli assi. La formula ottenuta è dunque proprio quella di Eulero: $k = k_1 \cos^2 \phi + k_2 \sin^2 \phi$. Dalla stessa formula segue che le due curvature principali sono i valori estremo massimo e minimo.

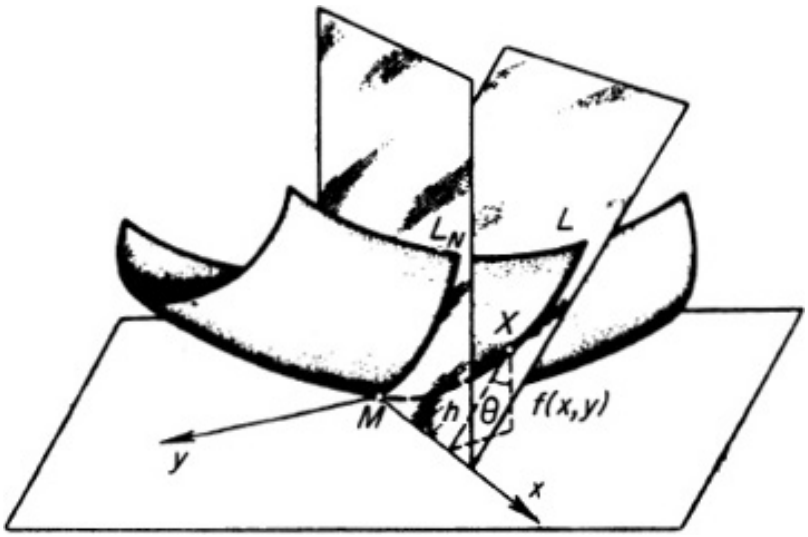


Figura 28.

2. Per dimostrare il teorema di Meusnier consideriamo una sezione normale L_N e una sezione L il cui piano formi un angolo θ con quello di L_N, come nella figura 28. Gli assi Ox e Oy giacciono nel piano tangente; scegliamo inoltre Ox tangente alle due curve nell'origine. La distanza $h(x, y)$ da Ox di un punto X di L, con coordinate $x, y, f(x, y)$, è evidentemente $|f(x, y)|/\cos\theta$ (figura 28), prendendo l'angolo non ottuso e $\cos\theta > 0$. Con la formula di Taylor esprimiamo la curvatura k_L così:

$$k_L = \lim_{x\to 0} \frac{2h(x,y)}{x^2} = \lim_{x\to 0} 2\,\frac{|f(x,y)|}{x^2\cos\theta}$$
$$= \lim_{x\to 0} \frac{|f_{xx}x^2 + 2f_{xy}xy + f_{yy}y^2 + 2\epsilon(x^2+y^2)|}{x^2\cos\theta}, \qquad (5)$$

dove $\epsilon \to 0$ per $x, y \to 0$. Poiché Ox è tangente a L, abbiamo $\lim_{x\to 0} y/x = 0$. Passando al limite nella (5), otteniamo

$$k_L = \frac{|f_{xx}|}{\cos\theta}.$$

Nel sistema scelto, però, L_N ha equazione $z = f(x, 0)$, quindi $|k_N| = |f_{xx}|$. Ne segue $k_L = |k_N|/\cos\theta$, e il teorema è dimostrato.

Curvatura media.

In molti problemi della teoria delle superfici, il ruolo principale non è svolto direttamente dalle curvature principali, ma da certe quantità che ne dipendono: la curvatura media e la curvatura gaussiana, detta anche totale, in un punto. Esaminiamole.

La curvatura media in un punto è la media aritmetica delle curvature principali:

$$K_{av} = \tfrac{1}{2}(k_1 + k_2).$$

Per mostrare l'utilità di questo concetto, consideriamo un problema meccanico. Sulla superficie di un corpo *F* sia tesa una membrana elastica di gomma. Qual è la pressione che essa esercita in ogni punto della superficie?

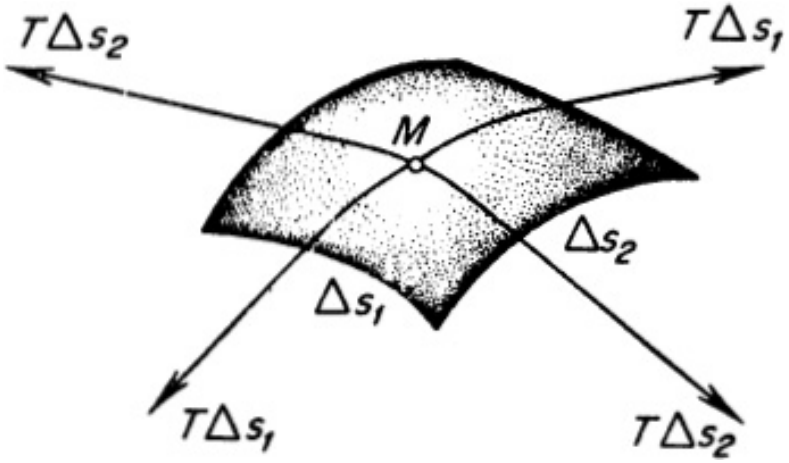


Figura 29.

La pressione in un punto *M* è misurata dalla forza per unità di area esercitata dalla membrana su una piccola porzione contenente *M*; più precisamente è il limite del rapporto fra tale forza e l'area della porzione, quando questa si restringe al punto *M*.

Circondiamo *M* con un piccolo rettangolo curvilineo, i cui lati di lunghezze Δs_1 e Δs_2 siano perpendicolari rispettivamente alla prima e alla seconda direzione principale in *M* (figura 29).[16] Su ogni lato agisce una forza proporzionale alla sua lunghezza e alla tensione uniforme *T* della membrana. Sui lati perpendicolari alla prima direzione principale agiscono quindi forze approssimativamente uguali a $T\Delta s_1$, tangenti alla superficie. Sull'altra coppia di lati agiscono forze di intensità $T\Delta s_2$. Per trovare la pressione in *M*, dividiamo la risultante delle quattro forze per l'area del rettangolo, approssimativamente $\Delta s_1 \Delta s_2$, e passiamo al limite per $\Delta s_1, \Delta s_2 \to 0$. Cominciamo dalle prime due forze.

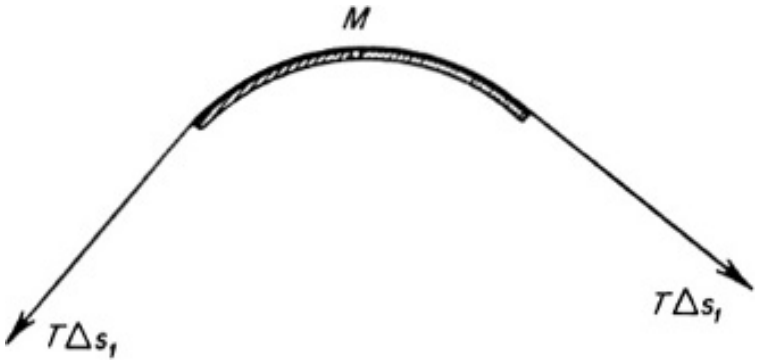


Figura 30.

Guardando il rettangolo di lato (figura 30), vediamo che queste forze sono dirette lungo le tangenti alla prima sezione normale e che i loro punti d'applicazione distano circa Δs_2. Ritroviamo quindi il problema del §2 sulla pressione di una fune. Usando quel risultato, il

limite richiesto è $k_1 T$, dove k_1 è la curvatura della prima sezione normale. Considerando nello stesso modo le altre due forze, otteniamo

$$P_M = T(k_1 + k_2) = 2TK_{\text{av}} .$$

Questo risultato ha molte conseguenze importanti. Vediamone un esempio.

È noto che lo strato superficiale di un liquido è sottoposto a una tensione uguale in tutte le direzioni. Per una massa liquida delimitata da una superficie curva, tale tensione produce, secondo il risultato precedente, una pressione proporzionale alla curvatura media nel punto considerato.

Nelle gocce di diametro molto piccolo la pressione è quindi molto grande, e ciò ne ostacola la formazione. In un vapore che si raffredda, le gocce si formano di regola intorno a granelli di polvere e particelle cariche. In un vapore perfettamente puro e moderatamente sovrasaturo la condensazione è ritardata. Se però una particella veloce attraversa il vapore ionizzandone le molecole, attorno agli ioni prodotti lungo il cammino compaiono per breve tempo piccole gocce, che rendono visibile la traccia della particella. Su questo principio si basa la camera di Wilson, largamente usata in fisica nucleare per osservare il moto delle particelle cariche.

Poiché la pressione di un liquido in equilibrio è isotropa, in assenza di altre cause di variazione della pressione una goccia deve assumere una forma con curvatura media costante sulla superficie. Nell'esperimento di Plateau si prendono due liquidi di uguale peso specifico, così che una massa dell'uno rimanga sospesa in equilibrio nell'altro. Si può supporre che la forma del liquido sospeso dipenda soltanto dalla tensione superficiale,[17] e si osserva che esso assume la forma di una sfera. Il risultato suggerisce che ogni superficie chiusa, connessa e senza autointersezioni, di curvatura media costante sia una sfera. Il teorema è effettivamente vero nelle consuete ipotesi di regolarità, benché la dimostrazione rigorosa sia difficile.

Possiamo affrontare la questione anche da un altro lato. La tensione superficiale tende a ridurre l'area mentre il volume del liquido non cambia; è quindi naturale attendersi che la massa sospesa abbia la minima superficie per il volume assegnato. Si dimostra che il corpo con questa proprietà è la sfera.

Il legame fra pressione normale della membrana e curvatura media permette anche di determinare la forma di una pellicola di sapone tesa su un contorno. Se la pressione dell'aria è uguale sui due lati, la risultante normale delle tensioni non è equilibrata da alcuna reazione di un sostegno, che qui manca, e deve quindi annullarsi. Per la superficie cercata abbiamo pertanto

$$K_{\text{av}} = 0. \qquad (6)$$

Dall'espressione analitica della curvatura media si ottiene un'equazione differenziale; bisogna risolverla imponendo che la superficie passi per il contorno assegnato.[18] Questo difficile problema è stato oggetto di molte ricerche.

La stessa equazione (6) nasce dal problema di trovare la superficie di area minima delimitata da un dato contorno. Dal punto di vista fisico il legame è naturale: la pellicola tende a ridurre

la propria area e raggiunge un equilibrio stabile quando l'area è minima rispetto alle variazioni ammesse. Per il loro rapporto con questo problema, le superfici di curvatura media nulla si chiamano minime, anche se l'equazione da sola esprime una condizione di stazionarietà e non garantisce un minimo.

Lo studio matematico delle superfici minime è molto interessante, anche per la grande varietà di forme essenzialmente diverse rivelata dagli esperimenti con le pellicole di sapone. La figura 31 mostra due pellicole tese su contorni differenti.

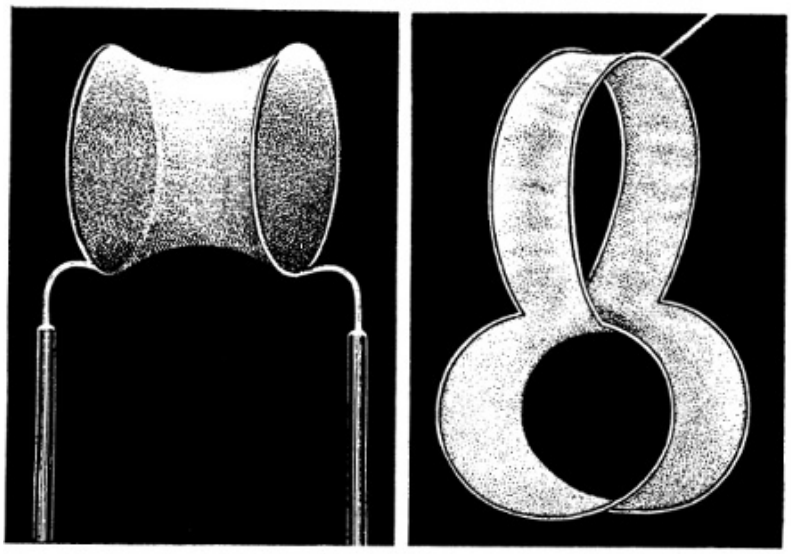


Figura 31.

Curvatura gaussiana.

La curvatura gaussiana di una superficie in un punto è il prodotto delle curvature principali:

$$K = k_1 k_2 .$$

Il suo segno determina il carattere della superficie vicino al punto. Per $K > 0$ la superficie ha forma di coppa, perché k_1 e k_2 hanno lo stesso segno; per $K < 0$ ha forma di sella. Gli altri casi già discussi corrispondono a curvatura gaussiana nulla. Il valore assoluto di K misura un aspetto complessivo dell'incurvamento della superficie, astraendo dalle singole direzioni. Ciò diventa particolarmente chiaro con una definizione alternativa che non richiede lo studio delle curve sulla superficie.

Prendiamo una piccola porzione G della superficie F, contenente M al suo interno, e costruiamo la normale in ogni punto.

Traslando tutte le origini delle normali in uno stesso punto, esse individuano un angolo solido (figura 32). La sua ampiezza dipende dall'area di G e dall'incurvamento della superficie in quella porzione. Possiamo quindi caratterizzare l'incurvamento mediante il rapporto fra l'ampiezza dell'angolo solido e l'area di G; è naturale definire una misura locale come limite di questo rapporto quando G si restringe a M.[19] Con l'opportuna interpretazione dell'area dell'immagine delle normali, tale limite è il valore assoluto della curvatura gaussiana in M.

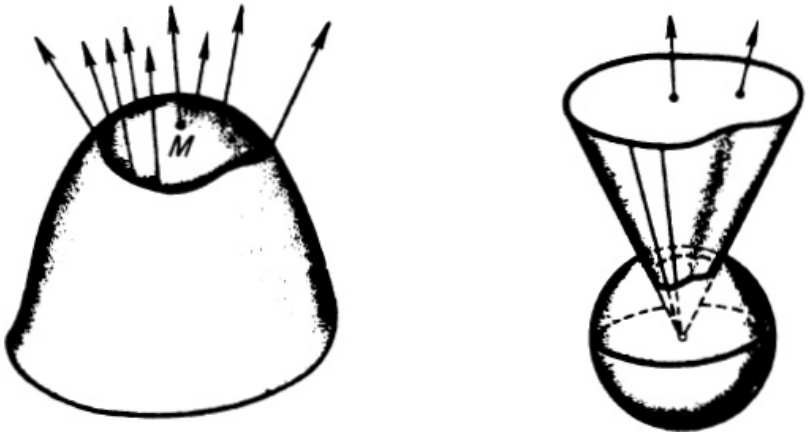


Figura 32.

La proprietà più notevole della curvatura gaussiana, che ne spiega l'importanza, è la seguente. Immaginiamo una superficie modellata in un materiale flessibile ma inestensibile, per esempio una lamina sottilissima, che possiamo piegare senza stirarla né strapparla. Le curvature principali cambiano durante la deformazione, ma Gauss dimostrò che il loro prodotto $k_1 k_2$ rimane invariato in ciascun punto. Due superfici con curvature gaussiane diverse nei punti corrispondenti sono dunque intrinsecamente diverse: non possono essere sovrapposte mediante deformazioni che conservino tutte le lunghezze. Per esempio, una porzione di sfera non può essere applicata senza stiramento su un piano o su una sfera di raggio diverso.

Abbiamo esaminato alcuni concetti fondamentali della teoria delle superfici. I metodi consistono principalmente, come già detto, nell'applicazione dell'analisi e soprattutto delle equazioni differenziali. Le dimostrazioni dei teoremi di Eulero e di Meusnier ne danno esempi semplici. Per problemi più complessi occorre un procedimento specifico che traduca le questioni geometriche in questioni analitiche. Esso si basa sulle cosiddette coordinate curvilinee e fu usato estesamente per la prima volta da Gauss nei problemi del tipo che discuteremo nel prossimo paragrafo.

§4. Geometria intrinseca e deformazione delle superfici

Geometria intrinseca.

Come già indicato, qui per deformazione di una superficie intendiamo un cambiamento di forma che conservi la lunghezza di tutte le curve tracciate su di essa. Arrotolare un foglio di carta in un tubo cilindrico costituisce, dal punto di vista geometrico, una deformazione di una porzione di piano: la carta non subisce praticamente alcuno stiramento e la lunghezza di ogni curva disegnata su di essa rimane invariata. Si conservano anche altre grandezze geometriche, per esempio l'area delle figure. Le proprietà della superficie che dipendono soltanto dalla sua metrica e non cambiano sotto tali deformazioni costituiscono la cosiddetta geometria intrinseca.

Quali sono queste proprietà? La deformazione conserva le lunghezze delle curve, dunque tutte le proprietà che in ultima analisi si determinano mediante tali lunghezze, ossia con misure eseguite sulla superficie stessa, restano invariate. La geometria intrinseca si chiama perciò anche geometria sulla superficie. L'aggettivo «intrinseca» significa che essa studia le proprietà della superficie in sé, indipendentemente dal modo in cui è collocata nello spazio circostante.[20] Se, per esempio, uniamo due punti di un foglio con un segmento e poi pieghiamo il foglio (figura 33), il segmento diventa una curva, ma conserva la proprietà di essere il cammino più breve fra i punti sulla superficie. Questa proprietà appartiene quindi alla geometria intrinseca. La curvatura spaziale della linea, invece, dipende da come il foglio è stato piegato e non è intrinseca.

In generale, poiché le dimostrazioni della geometria piana non ricorrono alle proprietà dello spazio circostante, i suoi teoremi appartengono alla geometria intrinseca di ogni superficie ottenuta deformando il piano, entro i limiti della corrispondenza considerata. Si può dire che la geometria piana è la geometria intrinseca del piano.

Un altro esempio familiare è la geometria sulla sfera, con cui abbiamo abitualmente a che fare nelle misure sulla superficie terrestre. Esso illustra particolarmente bene la natura della geometria intrinseca. A causa del grande raggio della Terra, ogni porzione direttamente visibile ci appare piana; le differenze dalla geometria piana osservate misurando grandi distanze si presentano quindi come leggi proprie della «geometria terrestre», espressione delle proprietà della superficie stessa, più che come effetti della sua curvatura nello spazio.

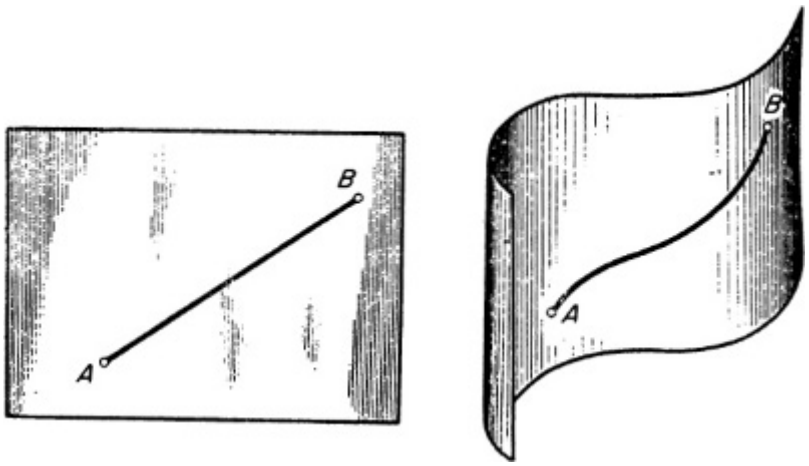


Figura 33.

Ricordiamo che Gauss giunse all'idea di studiare la geometria intrinseca attraverso i problemi della geodesia e della cartografia. Entrambe queste scienze applicate riguardano essenzialmente la geometria intrinseca della superficie terrestre. La cartografia studia in particolare le alterazioni dei rapporti fra le distanze quando una porzione della Terra viene rappresentata sul piano, e quindi la differenza fra geometria piana e geometria intrinseca terrestre.

Possiamo immaginare nello stesso modo la geometria intrinseca di una superficie qualsiasi. Supponiamo che vi vivano esseri così piccoli che, entro il loro campo visivo, la superficie appaia piana: sappiamo infatti che una porzione sufficientemente piccola di una superficie liscia differisce pochissimo dal piano tangente. Questi esseri non si accorgerebbero della curvatura nello spazio, ma misurando grandi distanze scoprirebbero leggi diverse da quelle del piano, corrispondenti alla geometria intrinseca della loro superficie. Che tali leggi cambino da una superficie all'altra si vede facilmente. Fissiamo un punto O e consideriamo una curva L i cui punti abbiano tutti distanza r da O, misurata sulla superficie lungo il cammino più breve (figura 34). Dal punto di vista intrinseco, L è una circonferenza di raggio r. La formula che ne esprime la lunghezza $s(r)$ in funzione di r appartiene alla geometria intrinseca. Essa cambia però con la superficie: sul piano $s(r) = 2\pi r$; sulla sfera di raggio R, come si verifica facilmente, $s(r) = 2\pi R \sin(r/R)$, per $0 < r < \pi R$. Sulla superficie della figura 35, a partire da un certo valore di r, la lunghezza della circonferenza con centro O dapprima rimane costante e poi diminuisce. Queste superfici hanno dunque geometrie intrinseche diverse.

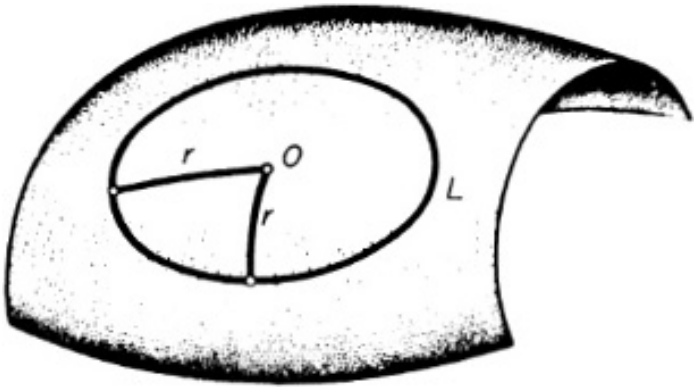


Figura 34.

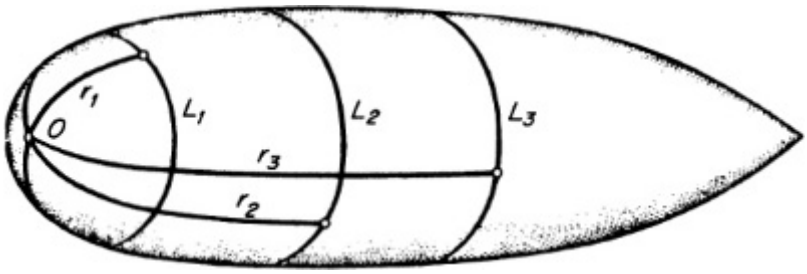


Figura 35.

Concetti fondamentali della geometria intrinseca.

Per illustrare la varietà dei suoi concetti e teoremi, torniamo alla geometria piana. Essa studia figure piane e proprietà solitamente espresse come relazioni fra grandezze fondamentali: lunghezza, angolo e area. Per dimostrare rigorosamente che angolo e area appartengono alla geometria intrinseca del piano, bisogna mostrare che si possono esprimere mediante lunghezze. È proprio così: un angolo si calcola conoscendo le lunghezze dei lati di un triangolo che lo contiene; anche l'area del triangolo si calcola dai lati, e per quella di un poligono basta dividerlo in triangoli.

Considerando la geometria piana come geometria intrinseca del piano, non siamo obbligati a limitarci alle nozioni scolastiche. Possiamo svilupparla quanto vogliamo e studiare nuovi problemi, purché siano formulabili, in ultima analisi, in termini di lunghezze. Possiamo introdurre la lunghezza di una curva, l'area di una regione delimitata da curve e così via: sono tutte nozioni della geometria intrinseca del piano.

Gli stessi concetti si introducono sulla superficie arbitraria. La lunghezza delle curve è il concetto iniziale; definire angoli e aree è un po' più complicato. Se la geometria intrinseca differisce da quella piana, non possiamo usare direttamente le formule consuete per esprimere angoli e aree mediante lunghezze. Ma vicino a un punto la superficie differisce poco dal piano tangente. Più precisamente, nelle opportune ipotesi di regolarità, proiettando una piccola porzione sul piano tangente in M, la distanza intrinseca fra punti vicini differisce dalla distanza fra le proiezioni per una quantità di ordine superiore al secondo rispetto alla scala della porzione. Nel definire grandezze locali mediante limiti nei quali intervengono infinitesimi di ordine non superiore al secondo, possiamo quindi sostituire la superficie con la

sua proiezione sul piano tangente. Le grandezze così ottenute risultano intrinseche. Questa possibilità è alla base delle definizioni fondamentali della geometria intrinseca.

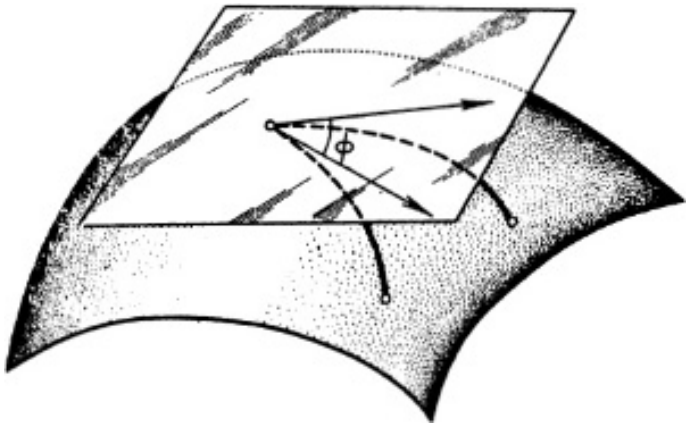


Figura 36.

Consideriamo, per esempio, angolo e area. Seguendo il principio generale, definiamo l'angolo fra due curve sulla superficie come quello fra le loro proiezioni sul piano tangente (figura 36). Esso coincide evidentemente con l'angolo fra le tangenti alle curve. Anche la definizione di area del §3 si basa sullo stesso principio. Infine, per descrivere l'incurvamento di una curva «all'interno» della superficie, introduciamo la «curvatura geodetica», nome che richiama le misure sulla Terra. In un punto essa è definita come la curvatura della proiezione della curva sul piano tangente (figura 37).

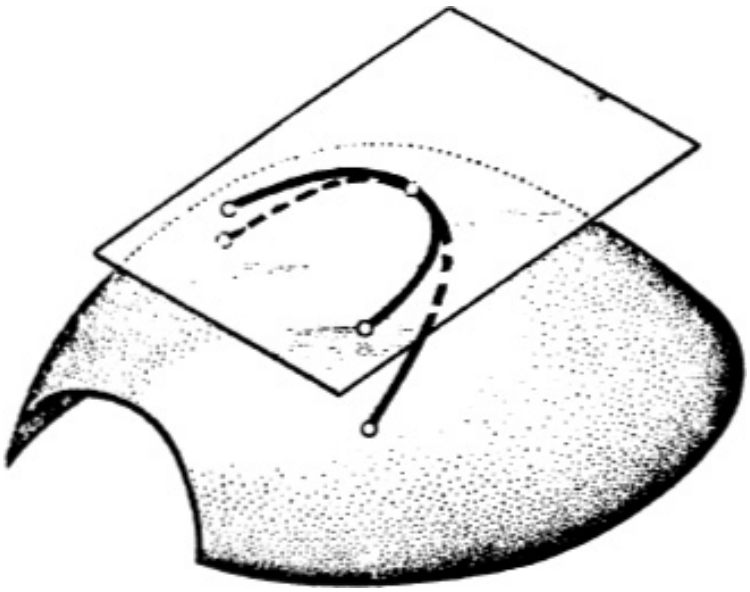


Figura 37.

I concetti fondamentali della geometria piana possono dunque essere introdotti nella geometria intrinseca di una superficie qualsiasi.

È facile definire anche figure analoghe a quelle piane. Abbiamo già considerato le circonferenze, definite come nel piano. Analogamente si definisce il corrispondente di un segmento: un segmento geodetico, ossia un cammino di lunghezza minima sulla superficie fra

due punti. È poi naturale chiamare triangolo una figura delimitata da tre segmenti geodetici, e così definire poligoni e altre figure. Poiché le loro proprietà dipendono dalla superficie, esistono in questo senso infinite geometrie intrinseche. Come ramo specifico della teoria delle superfici, la geometria intrinseca studia soprattutto le leggi generali valide per tutte e chiarisce come esse si esprimano mediante le grandezze caratteristiche di una data superficie.

Una delle caratteristiche più importanti, la curvatura gaussiana, non cambia nelle deformazioni e dipende quindi soltanto dalla geometria intrinseca. Essa misura inoltre, in modo notevole, quanto questa geometria differisca localmente da quella piana. Consideriamo una circonferenza L di raggio molto piccolo r, con centro in O. Sul piano la sua lunghezza è $s(r) = 2\pi r$. Su una superficie diversa dal piano, lo scarto da questo valore dipende essenzialmente, per piccoli r, dalla curvatura gaussiana K nel centro:

$$s(r) = 2\pi r - \frac{\pi}{3} K r^3 + \epsilon r^3,$$

dove $\epsilon \to 0$ per $r \to 0$. Per piccoli raggi possiamo dunque usare la formula piana trascurando termini del terzo ordine; l'errore, a meno di termini di ordine superiore al terzo, è proporzionale alla curvatura gaussiana. In particolare, se $K > 0$ la circonferenza piccola è più corta di quella piana di pari raggio; se $K < 0$ è più lunga. È facile farsene un'idea: vicino a un punto di curvatura positiva la superficie ha forma di coppa e le circonferenze risultano accorciate; vicino a un punto di curvatura negativa, la circonferenza sulla «sella» ha un andamento ondulato e risulta più lunga (figura 38).

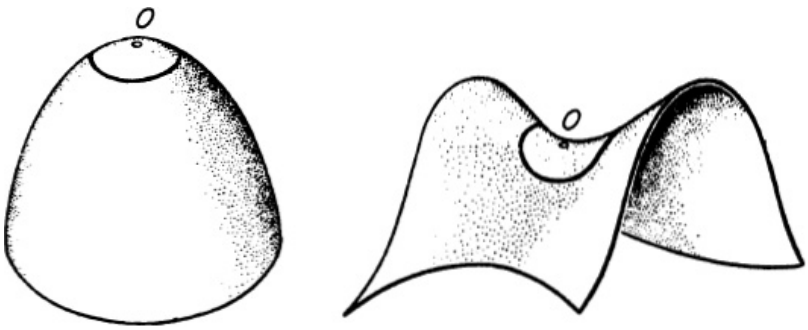


Figura 38.

Ne segue che una superficie con curvatura gaussiana variabile è geometricamente disomogenea: le proprietà intrinseche cambiano da punto a punto. La natura generale dei problemi avvicina la geometria intrinseca a quella piana, ma questa disomogeneità le distingue profondamente. Sul piano, per esempio, la somma degli angoli di un triangolo è pari a due angoli retti. Su una superficie arbitraria, invece, la somma degli angoli di un triangolo geodetico non è in generale determinata dal solo fatto che esso giaccia su una superficie nota e abbia lati di lunghezze assegnate. Conoscendo la curvatura gaussiana in ogni punto del triangolo, possiamo però calcolare la somma degli angoli α, β, γ mediante

$$\alpha + \beta + \gamma = \pi + \iint K \, d\sigma,$$

dove l'integrale è esteso alla regione triangolare. La formula contiene come casi particolari i noti teoremi sui triangoli del piano e della sfera unitaria. Nel primo caso $K = 0$ e $\alpha + \beta + \gamma = \pi$; nel secondo $K = 1$ e la somma è $\pi + S$, dove S è l'area del triangolo sferico.

Si può dimostrare che ogni porzione sufficientemente piccola di una superficie regolare con curvatura gaussiana identicamente nulla può essere deformata, o «sviluppata», sul piano, perché ha la stessa geometria intrinseca. Queste superfici si chiamano sviluppabili. Se la curvatura gaussiana è piccola, pur non potendo in generale sviluppare la superficie sul piano, su scale opportune la sua geometria intrinseca differisce poco da quella piana. Ancora una volta K misura lo scostamento dalla geometria del piano.

Linee geodetiche.

Nella geometria intrinseca il ruolo delle rette è svolto dalle linee geodetiche, o semplicemente «geodetiche».

Una retta nel piano può essere vista come una linea formata da segmenti che si sovrappongono parzialmente. Una geodetica si definisce nello stesso modo, sostituendo ai segmenti rettilinei quelli geodetici. In altri termini, è una curva sulla superficie ogni tratto sufficientemente piccolo della quale è un cammino minimo. Non ogni geodetica è minima globalmente: sulla sfera, per esempio, ogni arco di cerchio massimo è geodetico, ma è il cammino più breve fra gli estremi soltanto se non supera una semicirconferenza. Una geodetica può persino essere chiusa.

Per illustrare alcune proprietà importanti, consideriamo un modello meccanico.[21] Sulla superficie F sia tesa una fune elastica con estremi fissi (figura 39).[22] La fune è in equilibrio stabile quando ha lunghezza minima rispetto alle variazioni ammesse, perché ogni spostamento ne aumenterebbe la lunghezza e richiederebbe forze esterne. Localmente essa segue dunque una geodetica. Per l'equilibrio le forze elastiche devono essere compensate dalla reazione della superficie, diretta secondo la normale; supponiamo assenza di attrito. Nel §2 abbiamo visto che la forza dovuta alla tensione è diretta lungo la normale principale della curva seguita dalla fune. Otteniamo quindi questa proprietà: dove la curvatura della geodetica è non nulla, la sua normale principale è parallela alla normale alla superficie. Vale anche il viceversa: una curva regolare con questa proprietà è geodetica.

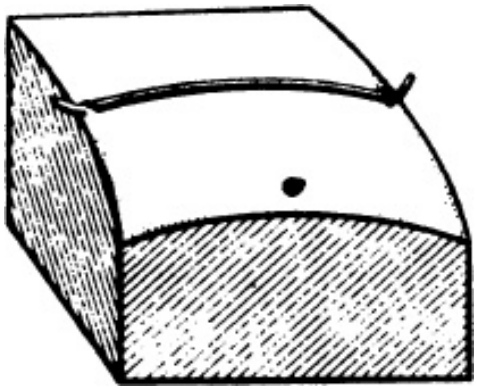


Figura 39.

Da qui segue un fatto importante: un punto materiale vincolato a muoversi sulla superficie, sul quale agisca soltanto la reazione normale del vincolo, percorre una geodetica. L'accelerazione normale è infatti diretta lungo la normale principale alla traiettoria; essendo la reazione della superficie l'unica forza, tale direzione coincide con quella della normale alla superficie. La traiettoria è dunque geodetica. Questo accresce l'analogia con le rette: come un punto libero si muove per inerzia lungo una retta, così un punto costretto a restare su una superficie, senza altre forze esterne, si muove lungo una geodetica.[23]

Dalla stessa proprietà segue che, se due superfici sono tangenti lungo una curva geodetica per una di esse, la curva è geodetica anche per l'altra. In ogni suo punto le superfici hanno infatti lo stesso piano tangente e quindi la stessa direzione normale. Questa coincide, dove definita, con la normale principale alla curva, che è pertanto geodetica su entrambe.

Ne derivano due proprietà intuitive. Anzitutto, se una striscia rettangolare flessibile e inestensibile, per esempio un righello d'acciaio, aderisce tangenzialmente a una superficie lungo la propria linea mediana, questa linea è una geodetica. Essa è infatti geodetica sulla striscia e quindi anche sulla superficie. Inoltre, facendo rotolare una superficie su un piano senza strisciamento né torsione, se il punto di contatto descrive una retta sul piano, la traccia corrispondente sulla superficie è una geodetica.[24] Entrambe le proprietà si osservano facilmente su un cilindro: la linea mediana di una striscia piana rettilinea appoggiata al cilindro (figura 40) coincide con una generatrice, una circonferenza oppure un'elica. Non è difficile dimostrare che le geodetiche di un cilindro circolare possono essere soltanto di questi tre tipi. Le stesse curve si ottengono facendo rotolare il cilindro su un piano sul quale sia tracciata una retta con il gesso.

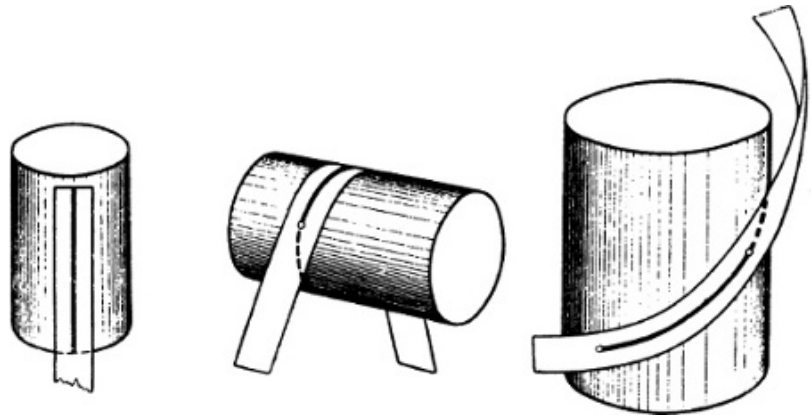


Figura 40.

L'analogia con le rette si completa con un'altra proprietà importante. Le rette del piano si possono definire come curve di curvatura nulla; le geodetiche di una superficie come curve di curvatura geodetica nulla. Ricordiamo che questa è la curvatura della proiezione sul piano tangente (figura 37). È naturale che la definizione coincida con la precedente: se tale proiezione ha curvatura nulla in ogni punto, la deviazione della curva dalla tangente avviene essenzialmente nella direzione normale alla superficie. Dove la normale principale è definita, essa segue dunque la normale alla superficie e la curva è geodetica. Viceversa, per una

geodetica la deviazione dalla tangente è normale alla superficie; proiettando sul piano tangente si ottiene una deviazione di ordine superiore e quindi curvatura nulla.

L'andamento delle geodetiche può essere molto diverso sulle varie superfici. La figura 41 ne mostra alcune su un iperboloide di rivoluzione.

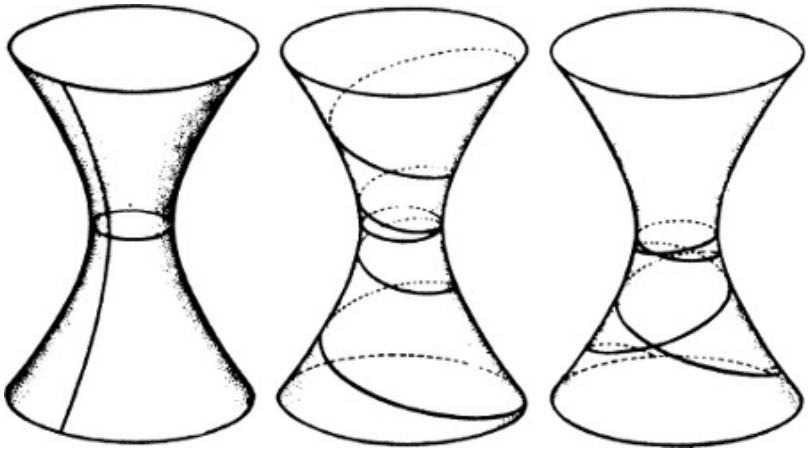


Figura 41.

Deformazione delle superfici.

Poiché la geometria intrinseca studia proprietà invarianti rispetto alle deformazioni, è naturale che studi anche le deformazioni stesse. La loro teoria è uno dei rami più interessanti e difficili della geometria e comprende molti problemi di enunciazione semplice che, all'epoca dell'opera, non erano ancora risolti.

Eulero e Minding avevano già considerato alcune questioni, ma i risultati generali per superfici arbitrarie arrivarono più tardi.

La prima domanda è se tutte le superfici ammettano deformazioni e in quale misura. Per le superfici analitiche, definite cioè da funzioni sviluppabili in serie di Taylor, Darboux risolse il problema alla fine dell'Ottocento. In particolare dimostrò il seguente risultato locale: scelta una geodetica sulla superficie e una curva analitica nello spazio della stessa lunghezza, con curvatura ovunque non nulla, una striscia sufficientemente stretta attorno alla geodetica può essere deformata in modo che questa coincida con la curva assegnata.[25] Una striscia di superficie può dunque essere deformata con notevole libertà. Si dimostra però che, imponendo l'immagine della geodetica come sopra, vi sono al più due possibili configurazioni locali. Se la curva assegnata è piana, esse sono immagini speculari rispetto al suo piano. Quando la curva assegnata è una retta, l'ultima affermazione non vale, come mostrano le deformazioni di una superficie cilindrica.

Abbiamo definito la deformazione come una trasformazione che conserva le lunghezze di tutte le curve. Finora ne abbiamo considerato soltanto il risultato finale, senza occuparci del processo intermedio. Ma, pensando a un materiale flessibile e inestensibile, è naturale considerare una trasformazione continua durante la quale le lunghezze si conservino in ogni istante. Queste trasformazioni si chiamano deformazioni continue.

A prima vista potrebbe sembrare che ogni deformazione si possa realizzare continuamente, ma non è così. È stato dimostrato, per esempio, che una superficie a forma di canale anulare (figura 42), nelle condizioni di regolarità considerate, non ammette deformazioni continue non banali. Ciò spiega anche perché un secchio con bordo ricurvo sia più robusto di uno con bordo semplice. Eppure sono possibili altre configurazioni isometriche: si può tagliare idealmente il canale lungo la circonferenza su cui poggia su un piano orizzontale e sostituire una metà con la sua immagine speculare rispetto al piano. Si confrontino le figure 42 e 43, nelle quali, per chiarezza, è disegnata soltanto la metà sinistra della superficie. Intuitivamente, l'ostacolo alla deformazione continua dipende dalla chiusura circolare del canale: per un canale rettilineo la stessa operazione può essere realizzata continuamente.

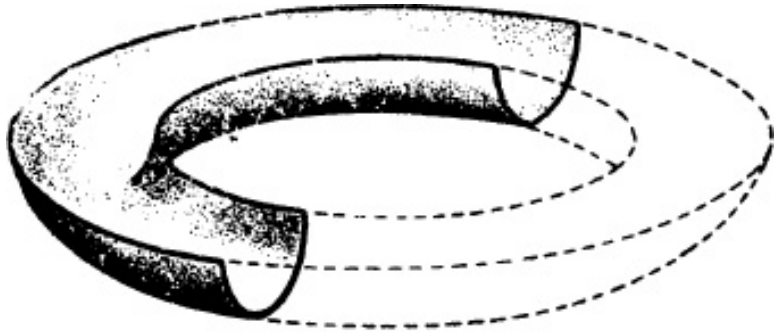


Figura 42.

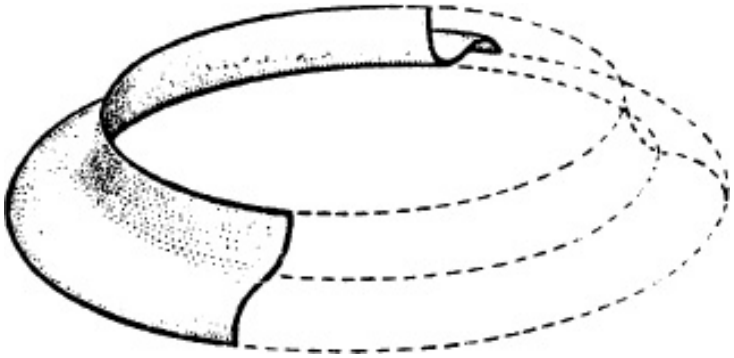


Figura 43.

Su una porzione sufficientemente piccola non sembrano esservi ostacoli evidenti alla deformazione continua. Ci si potrebbe quindi attendere che ogni deformazione locale sia realizzabile continuamente, eventualmente dopo una riflessione speculare. Il risultato vale nelle ipotesi considerate quando la curvatura gaussiana non si annulla nella porzione, nonché nel caso in cui sia identicamente nulla. Se invece si annulla in punti particolari, N. V. Efimov mostrò nel 1940 che porzioni arbitrariamente piccole di una superficie regolare possono non ammettere deformazioni continue non banali senza perdita della regolarità richiesta. Per esempio, la superficie $z = x^9 + \lambda x^6 y^3 + y^9$, con λ trascendente, ha la proprietà che nessuna porzione contenente l'origine, per quanto piccola, ammette deformazioni continue sufficientemente regolari. Il teorema di Efimov fu un risultato nuovo e piuttosto inatteso della geometria differenziale classica.

Oltre a questi problemi generali, molta attenzione viene dedicata a tipi particolari di deformazione.

Il legame fra geometria intrinseca e forma della superficie nello spazio.

Sappiamo già che alcune proprietà della superficie e delle figure su di essa sono determinate dalla geometria intrinseca, pur essendo strettamente collegate a proprietà dipendenti dalla sua collocazione nello spazio, dette «estrinseche». Le curvature principali sono estrinseche, ma il loro prodotto, la curvatura gaussiana, è intrinseco. Un altro esempio: affinché la normale principale di una curva sia parallela alla normale alla superficie, nei punti in cui è definita, occorre e basta che la curva abbia la proprietà intrinseca di essere geodetica.

La geometria intrinseca determina quindi soltanto in parte la forma spaziale della superficie.

Tale dipendenza può essere espressa analiticamente mediante equazioni che contengono sia grandezze intrinseche sia grandezze che descrivono come la superficie si incurva nello spazio. Una è la formula di Gauss che esprime la curvatura gaussiana in termini intrinseci; altre due sono le equazioni di Peterson e Codazzi, ricordate nel §1.

Le equazioni di Gauss, Peterson e Codazzi esprimono completamente il legame locale fra geometria intrinseca e incurvamento nello spazio, poiché racchiudono, almeno implicitamente, tutte le condizioni di compatibilità fra i due tipi di proprietà.

Poiché la geometria intrinseca non determina del tutto la forma spaziale, chiediamoci quali proprietà estrinseche occorra ancora assegnare. Risulta che, se due superfici hanno la stessa geometria intrinseca e, nei punti e nelle direzioni corrispondenti, le curvature normali hanno gli stessi valori con segno, le superfici sono congruenti, nelle ipotesi del teorema fondamentale: un movimento rigido porta l'una sull'altra. Peterson scoprì questo teorema circa quindici anni prima di Bonnet, al cui nome viene solitamente associato.

Gli strumenti analitici della teoria delle superfici.

L'applicazione sistematica dell'analisi ha portato alla costruzione di strumenti particolarmente adatti a questa teoria. Il passo decisivo fu compiuto da Gauss con le coordinate curvilinee. Esse generalizzano naturalmente le coordinate cartesiane del piano e sono strettamente legate alla geometria intrinseca, per la quale la rappresentazione $z = f(x, y)$ è scomoda: le coordinate spaziali dei punti cambiano quando la superficie viene deformata. Per evitare questo inconveniente si scelgono coordinate sulla superficie stessa. Ogni punto è individuato da due numeri u e v, che gli restano associati anche dopo la deformazione; le coordinate spaziali x, y, z saranno di volta in volta funzioni di u e v. Questi due numeri si chiamano coordinate curvilinee. Fissando una delle coordinate, per esempio v, e variando l'altra, si ottiene una curva coordinata. Le curve coordinate formano sulla superficie un reticolo curvilineo analogo a quello cartesiano. Il familiare sistema di longitudine e latitudine è proprio un sistema di coordinate curvilinee sulla sfera: il reticolo è formato da meridiani e paralleli[26] (figura 44). Per descrivere la posizione spaziale della superficie dobbiamo assegnare ogni punto in funzione di u e v, per esempio mediante il vettore $r = r(u, v)$ dall'origine fissa al punto, detto vettore posizione della superficie. Ciò equivale ad assegnare le sue tre componenti x, y, z come funzioni di u, v.[27] Per definire una curva sulla superficie basta dare u e v come funzioni di un parametro t; il vettore posizione del punto mobile è allora la funzione composta

$$\boldsymbol{r}[u(t), v(t)].$$

Per le funzioni vettoriali, derivata e differenziale si definiscono esattamente come per le funzioni scalari. Dalla definizione della derivata come limite di $\Delta r/\Delta t$ per $\Delta t \to 0$ segue subito che la derivata del vettore posizione di una curva è un vettore tangente (figura 45). Rimangono valide le proprietà fondamentali delle derivate ordinarie, fra cui la regola della funzione composta:



Figura 44.

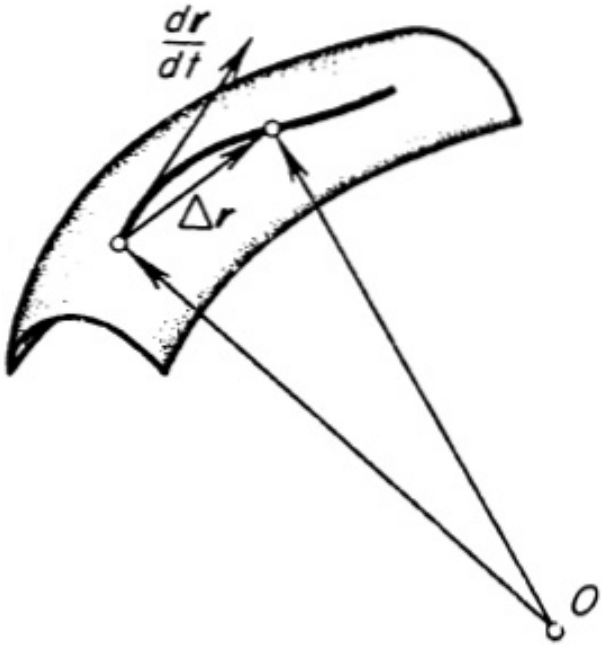


Figura 45.

$$\frac{d\boldsymbol{r}[u(t), v(t)]}{dt} = \frac{\partial \boldsymbol{r}}{\partial u}\frac{du}{dt} + \frac{\partial \boldsymbol{r}}{\partial v}\frac{dv}{dt} = \boldsymbol{r}_u u'_t + \boldsymbol{r}_v v'_t\,, \qquad (7)$$

dove r_u e r_v sono le derivate parziali della funzione vettoriale $r(u, v)$.

Si dimostra che la lunghezza di una curva è espressa dall'integrale

$$s = \int \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}\, dt.$$

Il differenziale della lunghezza d'arco è quindi

$$ds = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}\, dt.$$

Poiché $x'(t)$, $y'(t)$, $z'(t)$ sono le componenti del vettore $\boldsymbol{dr}/dt = \boldsymbol{r}'_t$, possiamo scrivere $ds = |\, \boldsymbol{r}'_t \,|\, dt$, dove $|\, \boldsymbol{r}'_t \,|$ denota il modulo del vettore $\boldsymbol{r}'_t$. Per le curve su una superficie, dalla (7) segue

$$ds = |\, \boldsymbol{r}_u u'_t + \boldsymbol{r}_v v'_t \,|\, dt.$$

Calcolando il quadrato del modulo del vettore a secondo membro con le regole dell'algebra vettoriale,[28] otteniamo

$$ds^2 = [\boldsymbol{r}_u^2 u_t'^2 + 2\boldsymbol{r}_u \boldsymbol{r}_v u'_t v'_t + \boldsymbol{r}_v^2 v_t'^2]\, dt^2\,.$$

Passando ai differenziali e introducendo le notazioni

$$\boldsymbol{r}_u^2 = E(u, v),\ \boldsymbol{r}_u \boldsymbol{r}_v = F(u, v),\ \boldsymbol{r}_v^2 = G(u, v),$$

abbiamo

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2.$$

Il quadrato del differenziale della lunghezza d'arco è dunque una forma quadratica nei differenziali *du*, *dv*, con coefficienti dipendenti dal punto. Si chiama prima forma quadratica fondamentale della superficie. Conoscendone i coefficienti *E*, *F*, *G* in ogni punto, possiamo calcolare la lunghezza di qualsiasi curva mediante

$$s = \int_{t_1}^{t_2} \sqrt{Eu_t'^2 + 2Fu_t'v_t' + Gv_t'^2}\,dt,$$

e la geometria intrinseca risulta così completamente determinata.

Mostriamo, per esempio, come esprimere angolo e area mediante *E*, *F*, *G*. Da un punto escano due curve, una data da $u = u_1(t)$, $v = v_1(t)$, l'altra da $u = u_2(t)$, $v = v_2(t)$. I rispettivi vettori tangenti sono

$$\boldsymbol{r}_1 = \boldsymbol{r}_u \frac{du_1}{dt} + \boldsymbol{r}_v \frac{dv_1}{dt},$$

$$\boldsymbol{r}_2 = \boldsymbol{r}_u \frac{du_2}{dt} + \boldsymbol{r}_v \frac{dv_2}{dt}.$$

Il coseno dell'angolo fra questi vettori è il loro prodotto scalare diviso per il prodotto dei moduli:

$$\cos\alpha = \frac{\boldsymbol{r}_1\boldsymbol{r}_2}{r_1 r_2} = \frac{\boldsymbol{r}_u^2 \frac{du_1}{dt}\frac{du_2}{dt} + \boldsymbol{r}_u\boldsymbol{r}_v\left(\frac{du_1}{dt}\frac{dv_2}{dt} + \frac{du_2}{dt}\frac{dv_1}{dt}\right) + \boldsymbol{r}_v^2 \frac{dv_1}{dt}\cdot\frac{dv_2}{dt}}{r_1 r_2}.$$

Ricordando che $\boldsymbol{r}_u^2 = E$, $r_u \cdot r_v = F$ e $\boldsymbol{r}_v^2 = G$, otteniamo

$$\cos\alpha = \frac{E\frac{du_1}{dt}\frac{du_2}{dt} + F\left(\frac{du_1}{dt}\frac{dv_2}{dt} + \frac{du_2}{dt}\frac{dv_1}{dt}\right) + G\frac{dv_1}{dt}\frac{dv_2}{dt}}{\sqrt{E\left(\frac{du_1}{dt}\right)^2 + 2F\frac{du_1}{dt}\frac{dv_1}{dt} + G\left(\frac{dv_1}{dt}\right)^2}\sqrt{E\left(\frac{du_2}{dt}\right)^2 + 2F\frac{du_2}{dt}\frac{dv_2}{dt} + G\left(\frac{dv_2}{dt}\right)^2}}.$$

Per l'area consideriamo un rettangolo curvilineo delimitato da $u = u_0$, $v = v_0$, $u = u_0 + \Delta u$, $v = v_0 + \Delta v$. Lo approssimiamo con il parallelogramma del piano tangente individuato dai vettori $r_u \Delta u$ e $r_v \Delta v$, tangenti alle curve coordinate (figura 46). La sua area è $\Delta S = |r_u|\,|r_v|\,\Delta u\,\Delta v \sin\phi$, dove ϕ è l'angolo fra i due vettori. Poiché $\sin\phi = \sqrt{1 - \cos^2\phi}$, segue $\Delta s = |\boldsymbol{r}_u|\,|\boldsymbol{r}_v|\,\Delta u\,\Delta v\sqrt{1 - \cos^2\phi} = \sqrt{\boldsymbol{r}_u^2\boldsymbol{r}_v^2 - |\boldsymbol{r}_u|^2\,|\boldsymbol{r}_v|^2\cos^2\phi}\,\Delta u\,\Delta v$. Ricordando $\boldsymbol{r}_u^2 = E$, $\boldsymbol{r}_v^2 = G$ e $|r_u|\,|r_v|\cos\phi = r_u \cdot r_v = F$, otteniamo $\Delta s = \sqrt{EG - F^2}\,\Delta u\,\Delta v$. Sommando

le aree dei parallelogrammi e passando al limite per Δu, $\Delta v \to 0$, ricaviamo la formula $S = \iint_D \sqrt{EG - F^2}\, du\, dv$, dove l'integrazione è estesa al dominio D delle variabili u, v che descrive la porzione di superficie.

Le coordinate curvilinee sono quindi molto comode per lo studio della geometria intrinseca.

Anche il modo in cui la superficie si incurva nello spazio può essere caratterizzato da una forma quadratica in du, dv. Se n è un versore normale alla superficie in M e Δr è l'incremento del vettore posizione a partire da quel punto, lo scostamento con segno h dal piano tangente (figura 47) è $n \cdot \Delta r$. Sviluppando Δr con la formula di Taylor, otteniamo

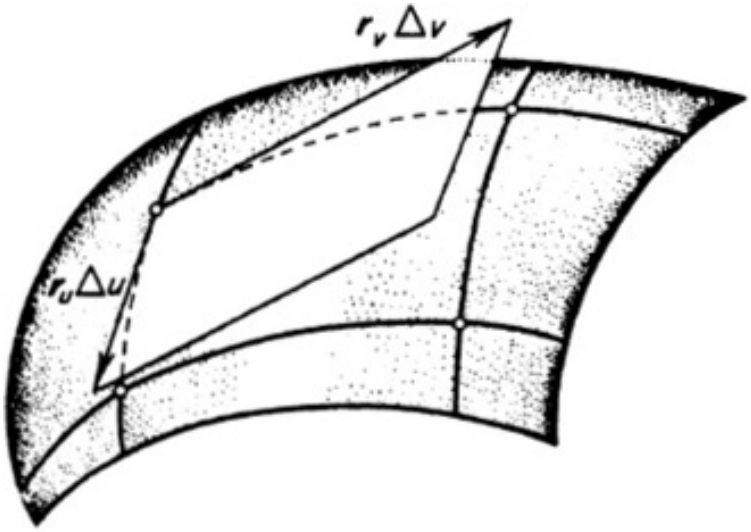


Figura 46.

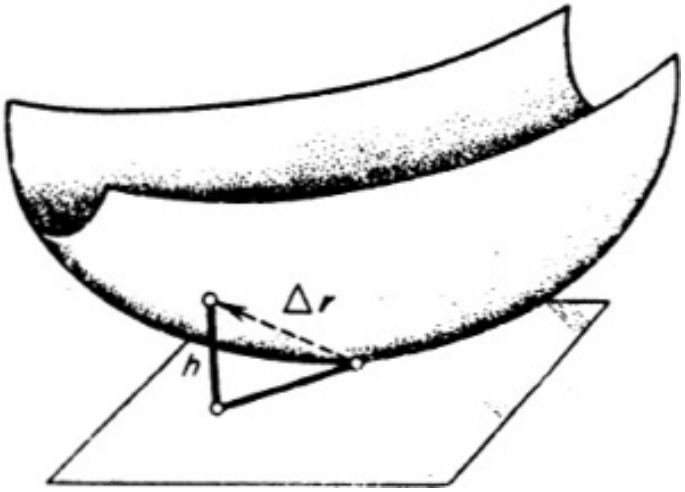


Figura 47.

$$h = n\, dr + \tfrac{1}{2} n\, d^2 r + \epsilon(du^2 + dv^2),$$

dove $\epsilon \to 0$ quando $\sqrt{du^2 + dv^2} \to 0$. Poiché dr giace nel piano tangente, $n \cdot dr = 0$. L'ultimo termine, $\epsilon(du^2 + dv^2)$, è trascurabile rispetto alla somma dei quadrati dei differenziali.

Rimane il termine principale $\frac{1}{2}\boldsymbol{n}\, d^2\boldsymbol{r}$. Il doppio della parte principale di h, cioè $n \cdot d^2 r$, è dunque una forma quadratica in du, dv:

$$\boldsymbol{n}\, d^2\boldsymbol{r} = \boldsymbol{n}\boldsymbol{r}_{uu}\, du^2 + 2\boldsymbol{n}\boldsymbol{r}_{uv}\, du\, dv + \boldsymbol{n}\boldsymbol{r}_{vv}\, dv^2.$$

Questa forma descrive come la superficie si allontana dal piano tangente. Si chiama seconda forma quadratica fondamentale. I suoi coefficienti, funzioni di u, v, vengono abitualmente indicati con

$$\boldsymbol{n}\boldsymbol{r}_{uu} = L, \quad \boldsymbol{n}\boldsymbol{r}_{uv} = M, \quad \boldsymbol{n}\boldsymbol{r}_{vv} = N.$$

Conoscendo la seconda forma fondamentale possiamo calcolare le curvature normali. Applicando $k = \lim_{l \to 0} 2h/l^2$, troviamo che la curvatura della sezione normale nella direzione corrispondente al rapporto du/dv è

$$k_n = \frac{\boldsymbol{n}\, d^2\boldsymbol{r}}{ds^2} = \frac{L\, du^2 + 2M\, du\, dv + N\, dv^2}{E\, du^2 + 2F\, du\, dv + G\, dv^2}.$$

Per una curva che non sia una sezione normale, il teorema di Meusnier lega questa curvatura a quella della curva mediante il coseno dell'angolo fra la normale principale della curva e la normale alla superficie. Quando il coseno è non nullo, si ricava la curvatura della curva dividendo opportunamente per esso.

La seconda forma fondamentale consente così uno studio analitico dell'incurvamento spaziale. In particolare, i teoremi di Eulero e Meusnier, le espressioni della curvatura gaussiana e media e altri risultati possono essere ricavati per via puramente analitica.

Il teorema di Peterson ricordato sopra mostra che le due forme quadratiche, insieme, determinano localmente una superficie a meno della sua posizione nello spazio. Lo studio analitico delle proprietà della superficie si riconduce quindi a quello delle due forme. I loro coefficienti non sono però indipendenti: il legame fra geometria intrinseca e incurvamento spaziale è espresso da tre relazioni, le equazioni di Gauss-Codazzi, fra i coefficienti della prima e della seconda forma fondamentale.

§5. Nuovi sviluppi della teoria delle curve e delle superfici

Famiglie di curve e di superfici.

Benché la teoria fondamentale fosse in gran parte compiuta già a metà Ottocento, essa ha continuato a svilupparsi in nuove direzioni, ampliando molto l'insieme delle figure e delle proprietà studiate dalla geometria differenziale. Una di queste direzioni risale in realtà agli inizi della disciplina: la teoria delle «famiglie», ossia degli insiemi continui di curve e superfici. La si può considerare nuova nel senso che i suoi aspetti più profondi vennero studiati soltanto dopo il pieno sviluppo della teoria fondamentale.

In generale, un insieme continuo di figure si chiama famiglia a n parametri se ogni figura è determinata dai valori di n parametri e tutte le grandezze che ne caratterizzano posizione, forma e altre proprietà dipendono almeno con continuità da essi. Da questo punto di vista una curva è una famiglia di punti a un parametro e una superficie una famiglia di punti a due parametri. L'insieme di tutte le circonferenze del piano è una famiglia di curve a tre parametri: le due coordinate del centro e il raggio.

Il problema più semplice della teoria consiste nel trovare il cosiddetto inviluppo della famiglia. Una superficie è l'inviluppo di una famiglia di superfici se in ciascun suo punto è tangente a una superficie della famiglia e in questo modo è tangente a tutte. Per esempio, l'inviluppo di sfere di uguale raggio con centri su una retta è un cilindro (figura 48); se i centri percorrono un piano, l'inviluppo è formato da due piani paralleli. Analogamente si definisce l'inviluppo di una famiglia di curve. La figura 49 rappresenta schematicamente getti d'acqua di una fontana lanciati sotto angoli diversi: in un piano formano una famiglia di curve approssimativamente paraboliche, il cui inviluppo si distingue come contorno generale dei getti. Non ogni famiglia ha un inviluppo: per esempio, non ne ha una famiglia di rette parallele. Esiste un semplice procedimento generale per cercare gli inviluppi; nel caso delle curve piane esso fu dato da Leibniz.

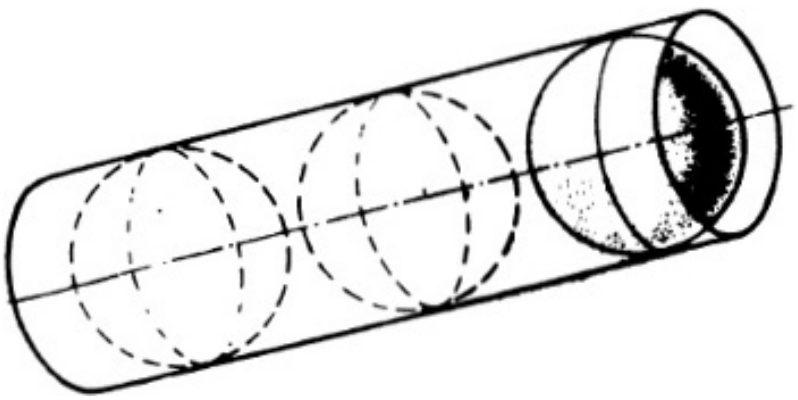


Figura 48.

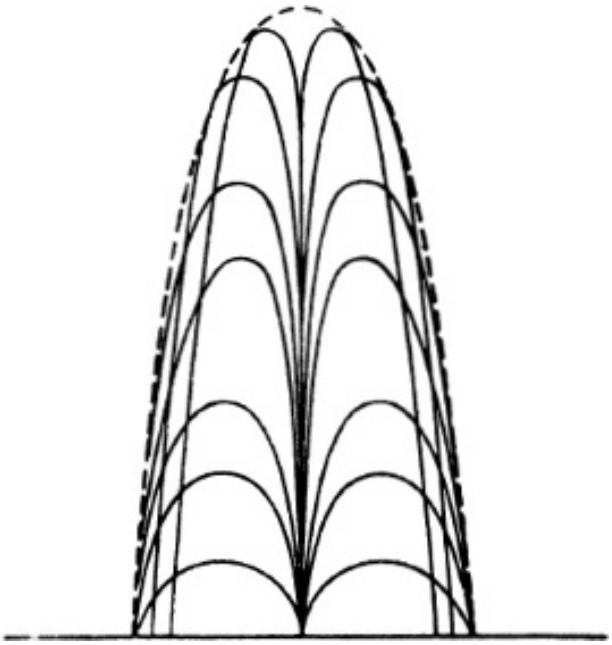


Figura 49.

Ogni curva regolare è, nel senso locale opportuno, l'inviluppo delle proprie tangenti, e ogni superficie quello dei propri piani tangenti. Questo offre un altro modo di definire curve e superfici, assegnando la famiglia delle tangenti o dei piani tangenti. In alcuni problemi è il modo più conveniente.

In generale i piani tangenti di una superficie sono diversi nei diversi punti e formano una famiglia a due parametri. In certi casi, come per un cilindro, la famiglia ha un solo parametro. Vale il notevole teorema seguente: una famiglia a un parametro di piani tangenti si presenta sulle superfici sviluppabili, cioè quelle di cui ogni porzione abbastanza piccola può essere deformata in una porzione piana, già ricordate nel §4. Localmente, nei casi analitici non degeneri, una superficie di questo tipo è formata da segmenti di rette: può essere cilindrica, con generatrici parallele, conica, con generatrici passanti per un punto, oppure formata dalle tangenti a una curva spaziale.

La teoria degli inviluppi è particolarmente utile nei problemi tecnici, per esempio nella teoria degli ingranaggi. Consideriamo due ruote dentate *A* e *B*. Per studiarne il moto relativo possiamo supporre fissa *A* e far muovere *B* intorno a essa (figura 50). Il profilo di un dente di *B*, nelle diverse posizioni, descrive una famiglia di curve nel piano di *A*; il profilo di *A* deve essere sempre tangente a queste curve, ossia deve esserne l'inviluppo. Non è una descrizione completa, poiché in un ingranaggio reale il contatto passa da una coppia di denti alla successiva, ma è la condizione fondamentale che ogni tipo di ingranaggio deve soddisfare.

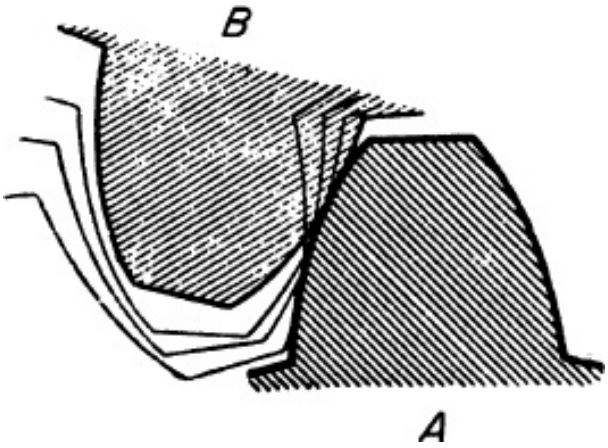


Figura 50.

Il problema degli inviluppi è relativamente semplice ed è stato risolto da molto tempo. La teoria delle famiglie di curve e superfici è però ricca di problemi interessanti quanto la teoria delle superfici stessa. Particolarmente sviluppata è la teoria delle «congruenze», cioè delle famiglie a due parametri di curve e, in particolare, di rette: le congruenze rettilinee. Vi si applicano sostanzialmente gli stessi metodi della teoria delle superfici.

La teoria delle congruenze rettilinee ebbe origine nella memoria di Monge «Sui trasporti di terra», dedicata agli sterri e ai riporti. Il titolo ne rivela lo scopo pratico: trovare il modo più conveniente di trasportare la terra da uno scavo a un rilevato.

Il suo sviluppo sistematico, iniziato a metà Ottocento, dipende in larga misura dal legame con l'ottica geometrica: un fascio di raggi luminosi in un mezzo omogeneo costituisce una congruenza rettilinea.

Superfici non regolari e geometria «in grande».

La teoria delle curve, delle superfici e delle loro famiglie, come si era costituita alla fine dell'Ottocento, si chiama abitualmente geometria differenziale classica e presenta le seguenti caratteristiche.

Anzitutto considera soltanto curve e superfici «sufficientemente lisce», cioè regolari, definite da funzioni dotate di un numero sufficiente di derivate. Superfici con cuspidi o spigoli, come quelle dei poliedri o il vertice di un cono, vengono quindi escluse oppure studiate soltanto nelle parti lisce.

In secondo luogo, la geometria differenziale classica si occupa soprattutto di porzioni abbastanza piccole di curve e superfici, cioè della geometria «in piccolo», lasciando in secondo piano le proprietà di un'intera superficie chiusa, proprie della geometria «in grande».

La deformazione delle superfici offre esempi tipici di questa distinzione. Già nel 1838 Minding mostrò che una porzione sufficientemente piccola di sfera può essere deformata: è un teorema «in piccolo». Nello stesso tempo congetturò che l'intera sfera non potesse esserlo. Il teorema, nelle condizioni classiche di regolarità, fu dimostrato da altri matematici soltanto nel 1899. È facile constatare sperimentalmente che un involucro sferico flessibile ma inestensibile resiste

alle deformazioni: una pallina da ping-pong conserva bene la forma, benché il materiale sia piuttosto flessibile. Un altro esempio è il secchio di lamiera del §4, rigido globalmente grazie al bordo ricurvo, mentre sue singole porzioni si piegano facilmente. Le proprietà «in piccolo» e «in grande» differiscono dunque in modo essenziale.

Altri esempi vengono dalla teoria delle geodetiche. Una geodetica «in piccolo», cioè su una piccola porzione di superficie, è un cammino minimo; «in grande» può non esserlo affatto e può persino essere chiusa, come i cerchi massimi della sfera.

I teoremi sulle geodetiche enunciati nel §4 sono essenzialmente locali. Appartengono invece alla geometria «in grande» le domande sul loro intero percorso. Su una superficie regolare, per esempio, due punti abbastanza vicini sono uniti da un'unica geodetica che rimanga interamente in un opportuno piccolo intorno. Se ammettiamo geodetiche che possano allontanarsene, un teorema presentato qui nella tradizione di Morse afferma che su una superficie chiusa due punti possono essere congiunti da infiniti segmenti geodetici, intesi come percorsi e contando quando necessario le diverse percorrenze. Per esempio, sul mantello di un cilindro circolare, due punti *A* e *B* possono essere uniti da eliche che compiono numeri diversi di giri. Anche il teorema sulle geodetiche chiuse proposto da Poincaré e dimostrato da Ljusternik e Šnirel'man, ricordato nel capitolo XVIII, §5, appartiene alla geometria «in grande».

Le dimostrazioni di questi teoremi, come di molti risultati globali, non erano accessibili con i soli strumenti usuali della geometria differenziale classica e richiesero nuovi metodi.

Quando i problemi globali attirarono inevitabilmente l'attenzione dei matematici, non fu più possibile mantenere la restrizione alle superfici regolari. Incontriamo continuamente superfici non regolari o con curvatura discontinua, come le lenti convesse con bordo tagliente. Inoltre molte superfici analitiche non si possono prolungare naturalmente senza incontrare singolarità, come spigoli o cuspidi, e diventare quindi non regolari.

Una porzione di cono, per esempio, non può essere prolungata naturalmente senza raggiungere il vertice, dove viene meno la regolarità.

È un caso particolare del seguente notevole risultato: prolungando naturalmente una sviluppabile analitica non cilindrica, si incontra in generale uno spigolo di regresso, o un vertice nel caso di un cono, oltre il quale non si può proseguire mantenendo la stessa regolarità.

Esiste dunque un legame profondo fra il comportamento globale della superficie e le sue singolarità. Per questo i problemi «in grande» e lo studio di spigoli, cuspidi, curvature discontinue e altre singolarità devono essere affrontati insieme.

Anche l'analisi seguì nuove direzioni analoghe. La teoria qualitativa delle equazioni differenziali, ricordata nel capitolo V, §7, studia le proprietà delle soluzioni nell'intero dominio di definizione, cioè «in grande», prestando particolare attenzione alle singolarità e ai punti singolari dell'equazione. L'analisi moderna studia inoltre funzioni non regolari estranee all'analisi classica (cfr. capitolo XV), offrendo alla geometria strumenti per superfici più

generali. Infine, nel calcolo delle variazioni, dove si cercano curve o superfici con proprietà estremali, può accadere che la figura limite sulla quale si raggiunge l'estremo non sia regolare. La classe considerata deve allora essere chiusa, cioè contenere le proprie figure limite: ciò conduce necessariamente allo studio almeno delle forme non regolari più semplici. Le nuove direzioni della geometria nacquero dunque in stretto collegamento con lo sviluppo complessivo della matematica.

L'interesse sistematico per i problemi globali e per le superfici non regolari cominciò circa cinquant'anni prima della stesura dell'opera e coinvolse molti matematici. Il primo passo essenziale fu compiuto da Hermann Minkowski (1864-1909), che pose le basi di un ampio ramo della geometria: la teoria dei corpi convessi. Fra i problemi che lo avviarono a queste ricerche vi era quello dei reticoli regolari, strettamente connesso alla teoria dei numeri e alla cristallografia geometrica.

Un corpo si dice convesso se per ogni punto del suo contorno passa un piano d'appoggio che non attraversa l'interno del corpo: esso può cioè essere appoggiato su un piano in ogni punto della superficie (figura 51). Un corpo convesso è determinato dalla propria superficie, sicché per molti scopi parlare di corpi convessi o di superfici convesse chiuse è equivalente. I teoremi generali vengono di solito dimostrati senza ipotesi ulteriori sulla regolarità della superficie e riguardano l'intero corpo. Le limitazioni classiche vengono così superate automaticamente. Tuttavia la teoria dei corpi convessi e quella delle superfici non regolari rimasero inizialmente poco collegate; la loro unificazione avvenne molto più tardi.

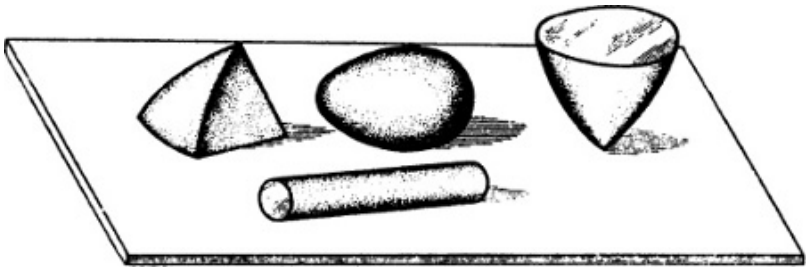


Figura 51.

A partire dal 1940 A. D. Aleksandrov sviluppò la teoria delle curve e delle superfici generali, comprendente sia le superfici regolari classiche sia forme non lisce come superfici poliedriche, superfici di corpi convessi arbitrari e altre ancora. Nonostante la grande generalità, la teoria si basa soprattutto su concetti e metodi geometrici intuitivi, pur utilizzando in modo essenziale l'analisi moderna. Uno dei metodi fondamentali consiste nell'approssimare le superfici mediante superfici poliedriche. Nella forma più semplice è un procedimento noto a scuola, per esempio nel calcolo dell'area laterale del cilindro come limite delle aree dei prismi. In molti casi produce risultati potenti, difficili da ricavare altrimenti o dimostrabili analiticamente soltanto introducendo idee complesse. L'essenza del metodo è ottenere prima il risultato per i poliedri e poi estenderlo alle superfici generali mediante un passaggio al limite.

Uno dei punti di partenza della teoria generale delle superfici convesse fu il teorema sulle condizioni che permettono di incollare uno sviluppo piano assegnato (figura 52) per ottenere un poliedro convesso. Di formulazione elementare, il teorema ha una dimostrazione non

elementare e conseguenze molto ampie. Il lettore conosce certamente la costruzione di superfici mediante incollaggio: per esempio il cubo ottenuto dallo sviluppo a croce della figura 52 o il cilindro da un rettangolo e due cerchi. Questa semplice operazione viene trasformata in un metodo generale di «taglio e incollaggio», che ha prodotto risultati profondi nella teoria delle superfici e ha trovato applicazioni pratiche.

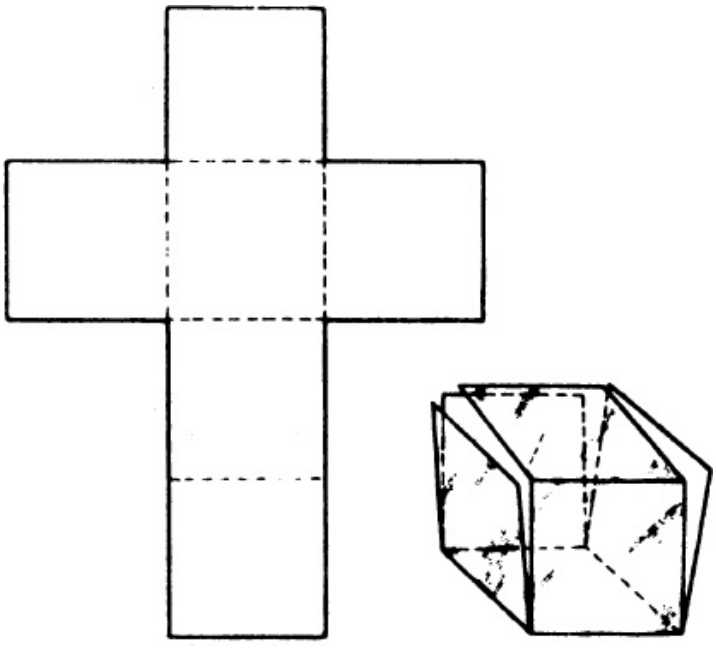


Figura 52.

A. V. Pogorelov ottenne risultati profondi in questa teoria. In particolare dimostrò la rigidità delle superfici convesse chiuse: due di esse con la stessa geometria intrinseca sono congruenti. Il risultato, raggiunto nel 1949, coronò gli sforzi di molti matematici che nei cinquant'anni precedenti erano riusciti a dimostrarlo soltanto con ipotesi aggiuntive. Insieme al metodo di incollaggio, i risultati di Pogorelov fornirono una soluzione completa di questo problema e chiarirono ampiamente la questione della deformabilità delle superfici convesse, chiuse e non chiuse. Stabilirono inoltre un legame stretto fra la nuova teoria e la geometria differenziale classica.

Si costruì così una teoria delle superfici che comprendeva quella classica, quella dei poliedri e delle superfici convesse arbitrarie, nonché classi molto ampie di superfici non convesse. Lo spazio non consente di discuterne in dettaglio risultati e problemi ancora aperti all'epoca, benché molti siano facilmente visualizzabili e non richiedano conoscenze speciali per essere compresi, a differenza delle difficili dimostrazioni.

Nel §4, parlando di deformazioni, intendevamo deformazioni di superfici regolari che ne conservassero la regolarità. Nel teorema di Pogorelov, invece, non si richiede regolarità né alla superficie iniziale né a quella finale: si impone però a entrambe la convessità.

È evidente che, per esempio, una sfera può essere deformata isometricamente se ammettiamo pieghe e perdita di convessità. Basta riflettere verso l'interno una calotta rispetto al piano del suo bordo, conservando l'incollaggio lungo quel bordo. Molto più inatteso fu il risultato, allora

recente, dell'americano Nash e dell'olandese Kuiper. Se conserviamo soltanto la regolarità del primo ordine e non imponiamo continuità, limitatezza o persino esistenza delle derivate seconde, diventano possibili deformazioni isometriche globali con grandissima libertà. In particolare, una superficie isometrica alla sfera può essere collocata dentro una palla arbitrariamente piccola, restando liscia nel senso del primo ordine e sviluppando increspature molto fini. Un'idea si ottiene pensando a un involucro sferico di tessuto morbidissimo, che si può spiegazzare in moltissime forme. Una pallina di celluloide si comporta invece diversamente: il materiale elastico resiste non soltanto allo stiramento, ma anche alle piegature brusche, e la rende molto rigida.

Geometria differenziale dei diversi gruppi di trasformazioni.

All'inizio del Novecento nacquero dalla geometria differenziale classica nuovi indirizzi fondati su un'idea comune: studiare le proprietà di curve, superfici e loro famiglie che restano invarianti sotto determinati tipi di trasformazioni. La teoria classica studia proprietà invarianti per movimenti rigidi; nulla vieta però di considerare altre trasformazioni geometriche. Una trasformazione proiettiva, per esempio, conserva le rette, e la geometria proiettiva studia le proprietà che non cambiano sotto tali trasformazioni. La geometria proiettiva ordinaria è affine, per i problemi trattati, alla geometria elementare e analitica; la geometria differenziale proiettiva, sviluppata all'inizio del Novecento, è invece analoga alla geometria differenziale classica, ma studia gli invarianti proiettivi di curve, superfici e famiglie. Fondamentali furono i contributi dell'americano Wilczynski, dell'italiano Fubini e del ceco Čech.

Nello stesso modo nacque la geometria differenziale affine, che studia proprietà invarianti sotto trasformazioni affini: trasformazioni che conservano non soltanto le rette, ma anche il parallelismo. Blaschke e i suoi allievi svilupparono questo ramo in una teoria generale. Ricordiamo inoltre la geometria conforme, che studia invarianti delle trasformazioni che conservano gli angoli fra le curve.

Le «geometrie» possibili sono molto varie: essenzialmente ogni gruppo di trasformazioni può essere assunto come base di una geometria, che studierà le proprietà delle figure lasciate invariate dalle trasformazioni del gruppo. Questo principio sarà approfondito nel capitolo XVII.

Altri nuovi indirizzi sono stati sviluppati con successo dai geometri sovietici S. P. Finikov, G. F. Laptev e altri. In questa panoramica non è però possibile rendere conto di tutte le ricerche nei diversi rami della geometria differenziale dell'epoca.

Letture consigliate

H. Busemann, Convex surfaces, Interscience, New York, 1958.

H. S. M. Coxeter, Introduction to geometry, Wiley, New York, 1961.

H. G. Eggleston, Convexity, Cambridge University Press, New York, 1958.

D. Hilbert e S. Cohn-Vossen, Geometry and the imagination, Chelsea, New York, 1952.

I. M. Yaglom e V. G. Boltyanskiĭ, Convex figures, Holt, Rinehart and Winston, New York, 1961.

Note al capitolo VII

[1] Le proprietà «in un punto» dipendono soltanto da un intorno arbitrariamente piccolo del punto. Sono definite mediante le derivate, calcolate in quel punto, delle funzioni che compaiono nelle equazioni della curva o della superficie. Per questo la geometria differenziale impone condizioni che garantiscano l'applicabilità del calcolo differenziale: le funzioni devono possedere un numero sufficiente di derivate.

[2] Gaspard Monge (1746-1818) fu non soltanto un grande scienziato, ma anche un attivo rivoluzionario francese: ministro della Marina e poi organizzatore della produzione di cannoni e polvere da sparo. Seguì il percorso, caratteristico di una parte della borghesia francese dell'epoca, dal giacobinismo all'adesione all'imperatore Napoleone.

[3] A rigore, per mantenere l'ordine di percorrenza è necessario che f sia monotona; una riparametrizzazione regolare richiede inoltre le opportune condizioni sulla derivata.

[4] Una curva spaziale può anche essere data come intersezione di due superfici, mediante il sistema $F(x, y, z) = 0$, $G(x, y, z) = 0$, nelle condizioni che assicurano una curva. Nelle trattazioni teoriche si usa spesso un vettore variabile: la posizione di X è individuata dal vettore $\mathbf{r} = \overrightarrow{OX}$, che va dall'origine al punto. Al variare di r, il suo estremo X percorre la curva (figura 3).

[5] L'esistenza di un limite finito, cioè della lunghezza della curva, non è evidente a priori, nemmeno per curve contenute in un dominio limitato. Se la curva è molto tortuosa, la sua lunghezza può essere grandissima; si possono persino costruire curve piane tanto «aggrovigliate» che nessun loro arco non banale abbia lunghezza finita, poiché le lunghezze delle spezzate inscritte crescono oltre ogni limite.

[6] La posizione limite della secante può non esistere, come nell'esempio della figura 13 del capitolo II. La curva $y = x \sin(1/x)$, completata con l'origine, oscilla vicino a zero in modo che la secante OA, per $A \to O$, continui a oscillare fra le rette OM e OL.

[7] Il risultato segue direttamente dalla definizione di tangente. Nella figura 7, $XX'/AX = \sin\alpha$, dove α è l'angolo fra tangente e secante. Per $\alpha \to 0$ anche il rapporto tende a zero.

[8] La curvatura può essere definita anche confrontando la curva con una circonferenza che funge da modello. Dove la curvatura è non nulla e valgono le opportune ipotesi, essa è il reciproco del raggio del cerchio osculatore, l'unico che approssima la curva al secondo ordine nel punto dato.

[9] Sarebbe più naturale scegliere un tratto che contenga A al suo interno. Il risultato non cambierebbe, ma il calcolo sarebbe un po' più complicato.

[10] Si dimostra che un'elica circolare ha torsione costante. Nei casi non degeneri si può quindi interpretare la torsione confrontando la curva con un'elica che ne riproduca localmente le caratteristiche. La torsione misura anche quanto una curva spaziale si discosti dall'essere piana: in analogia con la curvatura, caratterizza il suo allontanamento dal piano osculatore.

[11] È proprio l'espressione dell'area che verrà usata nel capitolo VIII, §1.

[12] Nel caso particolare $k_1 = k_2$, tutte le sezioni normali hanno la stessa curvatura, come accade sulla sfera.

[13] È facile verificare che $k(\phi) = k_1 \cos^2\phi + k_2 \sin^2\phi$ si annulla per $\phi = \arctan\sqrt{-k_1/k_2}$ e $\phi = \pi - \arctan\sqrt{-k_1/k_2}$, cambiando segno dapprima da positivo a negativo e poi da negativo a positivo, con i segni delle curvature scelti nel testo.

[14] Meusnier (1754-1793), matematico francese allievo di Monge, fu generale dell'esercito rivoluzionario e morì per ferite ricevute in battaglia.

[15] Indichiamo le derivate parziali con indici: per esempio f_x al posto di $\partial f/\partial x$, f_{yy} al posto di $\partial^2 f/\partial y^2$, e così via.

[16] Il ragionamento non è rigoroso nella forma esposta. Stimando gli errori introdotti, si può però trasformarlo in una dimostrazione rigorosa.

[17] Si può trascurare l'aumento della pressione con la profondità, poiché è uguale nei due liquidi, che hanno lo stesso peso specifico. Sulla superficie comune gli incrementi della pressione interna ed esterna dovuti alla profondità si compensano.

[18] Per una superficie data da $z = z(x, y)$, l'equazione (6) assume la forma

$$(1 + z_y'^2)z_{xx}'' - 2z_x'z_y'z_{xy}'' + (1 + z_x'^2)z_{yy}'' = 0.$$

[19] Per misurare l'angolo solido si costruisce una sfera di raggio unitario con centro nel suo vertice. L'area della regione sferica intercettata dall'angolo viene assunta come sua misura (figura 32).

[20] Le idee della geometria intrinseca hanno condotto a un'ampia generalizzazione del concetto matematico di spazio e hanno così svolto un ruolo molto importante nella fisica moderna; si veda il capitolo XVII.

[21] Come già osservato, il ragionamento non costituisce una dimostrazione rigorosa delle proprietà delle geodetiche: serve a illustrarne le più importanti.

[22] Una fune tesa non rimane necessariamente aderente a una superficie non convessa. Per evitare eccezioni conviene immaginare la superficie formata da due strati fra i quali passa la fune.

[23] Qui «esterne» indica tutte le forze diverse dalla reazione della superficie.

[24] Questa proposizione non differisce essenzialmente dalla precedente: il rotolamento senza strisciamento e senza torsione di una superficie su un piano equivale, in un senso preciso, allo svolgimento di una striscia piana lungo la superficie.

[25] Si esclude il caso in cui la curva assegnata abbia curvatura nulla, perché si dimostra facilmente che una geodetica di una superficie con curvatura gaussiana positiva non può essere trasformata in una retta.

[26] È significativo che le coordinate geografiche e le loro applicazioni pratiche fossero note molto prima dell'introduzione cartesiana delle coordinate piane.

[27] Gauss naturalmente non usava la notazione vettoriale, ma definiva separatamente x, y, z come funzioni di u, v. I vettori, introdotti attraverso i lavori di Hamilton e Grassmann, furono dapprima usati soprattutto in fisica e soltanto più tardi, nel Novecento, divennero uno strumento abituale della geometria analitica e differenziale.

[28] Il quadrato del modulo di un vettore è il suo prodotto scalare con se stesso. Per il prodotto scalare valgono le consuete regole di distributività (cfr. capitolo III, §9).

Nota editoriale

La traduzione conserva le cinque sezioni, le cinquantadue figure, le formule e le ventotto note del capitolo. Sono corretti refusi negli indici e nella terminologia. Le notizie sullo stato delle ricerche si riferiscono all'epoca dell'opera.

Sono esplicitate, dove necessarie, le ipotesi di regolarità e di non degenerazione. Nel teorema di Meusnier l'angolo è quello fra il piano della sezione e il piano della corrispondente sezione normale; il caso del piano osculatore tangente richiede un trattamento distinto. Le superfici minime sono definite dalla curvatura media nulla: questa condizione esprime stazionarietà dell'area e non garantisce da sola un minimo. Il teorema sulle superfici chiuse a curvatura media costante è riferito a superfici connesse, regolari e immerse senza autointersezioni.

Nel §4, per determinare una superficie a meno di un movimento rigido occorrono i valori delle curvature normali con il loro segno, non soltanto i segni. Nell'esempio delle coordinate sferiche è corretto un indice ripetuto. Il rotolamento usato per definire il trasporto parallelo è inteso senza strisciamento e senza torsione. Nel risultato di Nash e Kuiper si parla di una superficie isometrica contenuta in una palla arbitrariamente piccola, non di una sfera rotonda più piccola. Queste precisazioni rendono esplicite le condizioni matematiche degli enunciati.

CAPITOLO VIII

Il calcolo delle variazioni

V. I. Krylov

§1. Introduzione

Esempi di problemi variazionali.

Per descrivere più chiaramente l'insieme dei problemi studiati dal calcolo delle variazioni,[1] cominciamo con alcuni problemi particolari.

1. La curva di discesa più rapida.

Il problema della brachistocrona, ossia della curva di discesa più rapida, fu storicamente il primo problema nello sviluppo del calcolo delle variazioni.

Fra tutte le curve che congiungono i punti M_1 e M_2, si richiede di trovare quella lungo la quale un punto materiale, partendo da M_1 senza velocità iniziale e muovendosi sotto l'azione della gravità, giunga a M_2 nel minor tempo possibile.

Per risolvere il problema dobbiamo considerare tutte le curve ammissibili che uniscono M_1 e M_2. A ogni curva determinata l corrisponde un valore determinato T del tempo impiegato dal punto materiale per percorrerla in discesa. Il tempo T dipende dalla scelta di l; fra tutte le curve che congiungono M_1 e M_2 dobbiamo scegliere quella cui corrisponde il valore minimo di T.

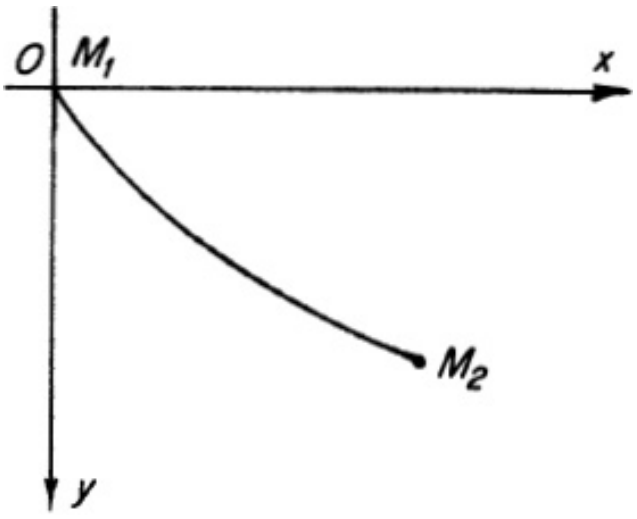


Figura 1.

Il problema della brachistocrona può essere formulato nel modo seguente.

Conduciamo un piano verticale per i punti M_1 e M_2. La curva di discesa più rapida giace in questo piano, e possiamo quindi limitarci a considerare curve piane. Prendiamo M_1 come origine, l'asse *Ox* orizzontale e l'asse *Oy* verticale, orientato verso il basso (figura 1). Le coordinate di M_1 sono (0, 0); indichiamo quelle di M_2 con (x_2, y_2). Consideriamo una curva arbitraria descritta dall'equazione

$$y = f(x), \quad 0 \leqslant x \leqslant x_2\,, \tag{1}$$

dove *f* è una funzione continuamente derivabile nell'interno dell'intervallo considerato. Poiché la curva passa per M_1 e M_2, agli estremi dell'intervallo $[0, x_2]$ essa deve soddisfare le condizioni

$$f(0) = 0, \quad f(x_2) = y_2\,. \tag{2}$$

Se prendiamo un punto arbitrario $M(x, y)$ sulla curva, la velocità v del punto materiale in quella posizione è legata alla coordinata y dalla ben nota relazione fisica

$$\tfrac{1}{2} v^2 = gy,$$

ossia

$$v = \sqrt{2gy}.$$

Il tempo necessario al punto materiale per percorrere un elemento d'arco *ds* della curva è

$$\frac{ds}{v} = \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}}\, dx,$$

e quindi il tempo totale di discesa lungo la curva, da M_1 a M_2, è uguale a

$$T = \frac{1}{\sqrt{2g}} \int_0^{x_2} \frac{\sqrt{1 + y'^2}}{\sqrt{y}}\, dx. \tag{3}$$

Trovare la brachistocrona equivale dunque a risolvere il seguente problema di minimo: fra tutte le funzioni ammissibili (1) che soddisfano le condizioni (2), trovare quella alla quale corrisponde il valore minimo dell'integrale (3).

2. La superficie di rotazione di area minima.

Fra le curve che congiungono due punti di un piano, si richiede di trovare quella il cui arco, ruotando intorno all'asse *Ox*, generi la superficie di area minima.

Indichiamo i punti dati con $M_1(x_1, y_1)$ e $M_2(x_2, y_2)$ e consideriamo una curva arbitraria di equazione

$$y = f(x). \tag{4}$$

Se la curva passa per M_1 e M_2, la funzione *f* soddisfa le condizioni

$$f(x_1) = y_1\,, \quad f(x_2) = y_2\,. \tag{5}$$

Ruotando intorno all'asse *Ox*, questa curva, supposta nel semipiano $y \geq 0$, descrive una superficie la cui area è numericamente uguale all'integrale

$$S = 2\pi \int_{x_1}^{x_2} y\,\sqrt{1 + y'^2}\,dx. \tag{6}$$

Il valore dipende dalla scelta della curva, ossia della funzione $y = f(x)$. Fra tutte le funzioni (4) che soddisfano le condizioni (5), dobbiamo trovare quella che rende minimo l'integrale (6).

3. Deformazione di una membrana omogenea.

Per membrana si intende di solito una superficie elastica che, nello stato di riposo, è piana, si flette liberamente e oppone resistenza soltanto all'estensione. Supponiamo che l'energia potenziale della membrana deformata sia proporzionale all'aumento della sua area.

Nello stato di riposo la membrana occupi un dominio *B* del piano *Oxy* (figura 2). Deformiamone il bordo in direzione perpendicolare a *Oxy* e indichiamo con $\phi(M)$ lo spostamento del punto *M* del bordo. Anche la parte interna della membrana si deforma; si richiede di trovarne la posizione di equilibrio per una data deformazione del bordo.

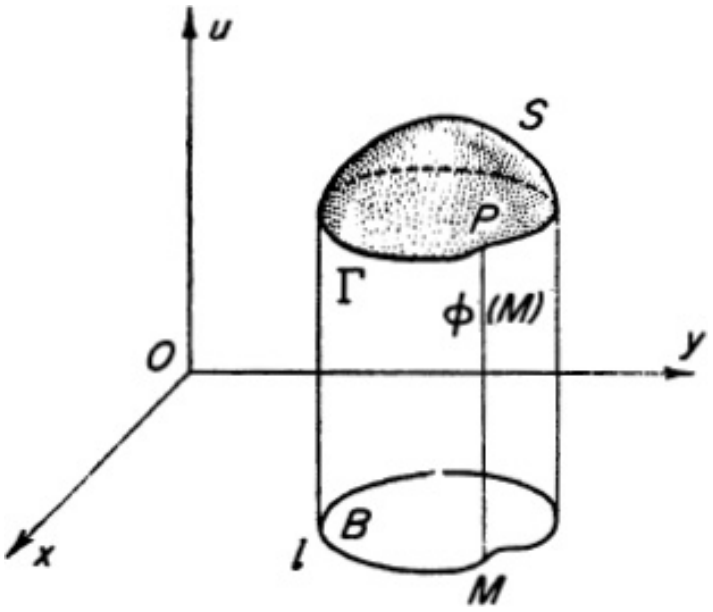


Figura 2.

Con buona approssimazione possiamo supporre che tutti i punti della membrana si spostino perpendicolarmente al piano *Oxy*. Indichiamo con $u(x, y)$ lo spostamento del punto (x, y). L'area della membrana nella posizione deformata è[2]

$$\iint_B (1 + u_x^2 + u_y^2)^{1/2}\,dx\,dy.$$

Se le deformazioni degli elementi della membrana sono abbastanza piccole da permettere di trascurare le potenze superiori delle derivate u_x e u_y, possiamo sostituire questa espressione dell'area con quella più semplice

$$\iint_B \left[1 + \frac{1}{2}(u_x^2 + u_y^2)\right] dx\, dy.$$

La variazione dell'area della membrana è uguale a

$$\frac{1}{2}\iint_B (u_x^2 + u_y^2)\, dx\, dy;$$

e quindi l'energia potenziale di deformazione vale

$$\frac{\mu}{2}\iint_B (u_x^2 + u_y^2)\, dx\, dy, \qquad (7)$$

dove μ è una costante che dipende dalle proprietà elastiche della membrana.

Poiché lo spostamento dei punti del bordo è assegnato, la funzione $u(x, y)$ deve soddisfare la condizione

$$u|_l = \phi(M) \qquad (8)$$

sul contorno del dominio B.

Nella posizione di equilibrio l'energia potenziale di deformazione deve assumere il valore minimo possibile. La funzione $u(x, y)$, che descrive lo spostamento dei punti della membrana, va dunque trovata risolvendo il seguente problema matematico: fra tutte le funzioni $u(x, y)$ continuamente derivabili nel dominio B e soddisfacenti la condizione (8) sul contorno, trovare quella che rende minimo l'integrale (7).

Estremi dei funzionali e calcolo delle variazioni.

Questi esempi danno un'idea dei problemi considerati; per precisare la posizione del calcolo delle variazioni nella matematica dobbiamo però introdurre alcuni concetti nuovi. Ricordiamo che uno dei concetti fondamentali dell'analisi matematica è quello di funzione. Nel caso più semplice, la dipendenza funzionale può essere descritta così. Sia M un insieme qualsiasi di numeri reali. Se a ogni numero x di M corrisponde un numero y, diciamo che sull'insieme M è definita una funzione $y = f(x)$. L'insieme M è spesso chiamato dominio di definizione della funzione.

Il concetto di funzionale è una generalizzazione diretta e naturale di quello di funzione, che comprende come caso particolare.

Sia M un insieme di oggetti di qualunque natura. Per ora non importa quali siano: possono essere numeri, punti di uno spazio, curve, funzioni, superfici, stati o anche moti di un sistema

meccanico. Per brevità li chiameremo elementi dell'insieme M e li indicheremo con la lettera x.

Se a ogni elemento x dell'insieme M corrisponde un numero y, diciamo che su M è definito un funzionale $y = F(x)$.

Se M è un insieme di numeri x, il funzionale $y = F(x)$ è una funzione di una variabile. Se M è un insieme di coppie di numeri (x_1, x_2), o di punti di un piano, il funzionale è una funzione $y = F(x_1, x_2)$ di due variabili, e così via.

Per il funzionale $y = F(x)$ poniamo il problema seguente:

Fra tutti gli elementi x di M, trovare quello per il quale il funzionale $y = F(x)$ assume il valore minimo.

Il problema del massimo di un funzionale si formula allo stesso modo.

Osserviamo che, cambiando il segno del funzionale $F(x)$ e considerando $-F(x)$, il massimo (minimo) di $F(x)$ diventa il minimo (massimo) di $-F(x)$. Non occorre quindi studiare separatamente massimi e minimi; nel seguito ci occuperemo soprattutto dei minimi dei funzionali.

Nel problema della curva di discesa più rapida, il funzionale di cui cerchiamo il minimo è l'integrale (3), cioè il tempo di discesa di un punto materiale lungo una curva. Esso è definito sulle funzioni ammissibili (1) che soddisfano le condizioni (2).

Nel problema dell'equilibrio di una membrana, il funzionale è l'energia potenziale (7) della membrana deformata, e ne cerchiamo il minimo nell'insieme delle funzioni $u(x, y)$ che soddisfano la condizione al contorno (8).

Ogni funzionale è definito da due elementi: l'insieme M degli oggetti x su cui è assegnato e la legge che associa a ciascun x un numero, il valore del funzionale. I metodi per cercarne i valori minimi e massimi dipendono certamente dalle proprietà dell'insieme M.

Il calcolo delle variazioni è un capitolo particolare della teoria dei funzionali. Esso considera funzionali definiti su insiemi di funzioni e si propone di costruire una teoria dei loro estremi.

Questo ramo della matematica acquistò particolare importanza quando se ne scoprirono i legami con numerosi problemi della fisica e della meccanica. La ragione è la seguente. Come vedremo, affinché una funzione renda estremo un funzionale è necessario che soddisfi una certa equazione differenziale. D'altra parte, come si è ricordato nei capitoli sulle equazioni differenziali, le leggi quantitative della meccanica e della fisica si esprimono spesso mediante equazioni di questo tipo. Si scoprì che molte di esse figuravano anche fra le equazioni differenziali del calcolo delle variazioni. Divenne così possibile interpretare equazioni della meccanica e della fisica come condizioni di estremo, o più generalmente di stazionarietà, per opportuni funzionali, ed esprimere le leggi fisiche richiedendo un valore estremo, in particolare un minimo, di determinate grandezze. Si introdussero così nuovi punti di vista: alcune leggi potevano essere sostituite da enunciati equivalenti in termini di «principi di

minimo». A sua volta, ciò aprì un nuovo metodo per risolvere esattamente o approssimativamente problemi fisici, cercando i minimi dei funzionali corrispondenti.

§2. Le equazioni differenziali del calcolo delle variazioni

L'equazione differenziale di Eulero.

Il lettore ricorderà che, affinché una funzione derivabile f abbia un estremo in un punto interno x, è necessario che in quel punto la derivata f' sia nulla: $f'(x) = 0$; equivalentemente, deve annullarsi il differenziale della funzione: $df = f'(x)\,dx = 0$.

Il nostro prossimo obiettivo è trovare l'analogo di questa condizione nel calcolo delle variazioni, ossia stabilire una condizione necessaria che una funzione deve soddisfare per rendere estremo un funzionale.

Mostreremo che tale funzione deve soddisfare una determinata equazione differenziale, la cui forma dipende dal funzionale considerato. Cominciamo dal cosiddetto integrale più semplice del calcolo delle variazioni, cioè da un funzionale rappresentato dall'integrale

$$I(y) = \int_{x_1}^{x_2} F(x, y, y')\,dx. \qquad (9)$$

La funzione F sotto il segno d'integrale dipende da tre argomenti (x, y, y'). Supporremo che sia definita e due volte continuamente derivabile rispetto a questi argomenti, per ogni valore di y' e per (x, y) in un certo dominio B del piano Oxy. Nel seguito si suppone di rimanere sempre nell'interno di tale dominio.

Naturalmente y è una funzione di x,

$$y = y(x), \qquad (10)$$

continuamente derivabile nell'intervallo $x_1 \leqslant x \leqslant x_2$, e y' è la sua derivata.

Geometricamente la funzione $y(x)$ è rappresentata nel piano Oxy da una curva l sopra l'intervallo $[x_1, x_2]$ (figura 3).

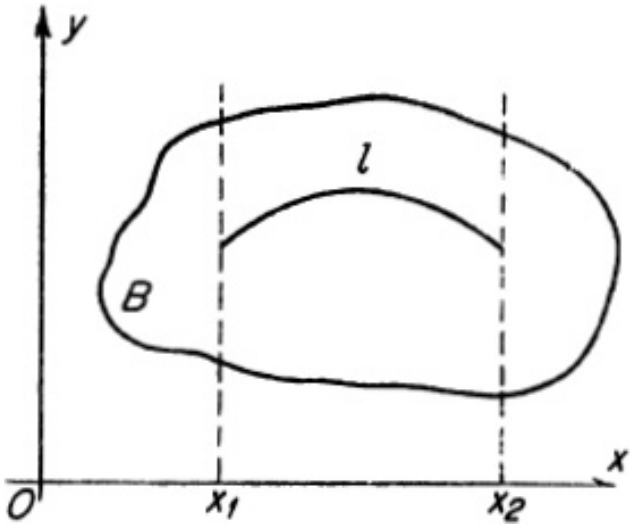


Figura 3.

L'integrale (9) generalizza gli integrali (3) e (6) incontrati nei problemi della curva di discesa più rapida e della superficie di rotazione di area minima. Il suo valore dipende dalla scelta della funzione $y(x)$, ossia della curva l, e il problema del suo minimo si interpreta così:

Dato un insieme M di funzioni (10), ossia di curve l, trovare fra esse la funzione, o curva, per la quale l'integrale $I(y)$ assume il valore minimo.

Dobbiamo innanzitutto precisare l'insieme M delle funzioni per le quali consideriamo l'integrale (9). Nel calcolo delle variazioni esse si chiamano di solito funzioni ammissibili al confronto. Consideriamo il problema con valori assegnati agli estremi. L'insieme delle funzioni ammissibili è definito dalle due condizioni seguenti:

1. $y(x)$ è continuamente derivabile nell'intervallo $[x_1, x_2]$;

2. Agli estremi dell'intervallo $y(x)$ assume valori assegnati,

$$y(x_1) = y_1\,, \quad y(x_2) = y_2\,. \tag{11}$$

Per il resto, $y(x)$ può essere del tutto arbitraria. Geometricamente consideriamo tutte le curve regolari sopra l'intervallo $[x_1, x_2]$, passanti per i punti $A(x_1, y_1)$ e $B(x_2, y_2)$ e rappresentabili mediante l'equazione (10). Supponiamo che esista una funzione che renda minimo l'integrale e indichiamola con $y(x)$.

Il semplice e ingegnoso ragionamento che segue, spesso utilizzabile nel calcolo delle variazioni, conduce a una forma particolarmente semplice della condizione necessaria cui deve soddisfare $y(x)$. In sostanza, esso riduce il problema del minimo dell'integrale (9) a quello del minimo di una funzione.

Consideriamo la famiglia di funzioni dipendente dal parametro numerico α,

$$\bar{y}(x) = y(x) + \alpha\eta(x). \tag{12}$$

Affinché $\bar{y}(x)$ sia ammissibile per ogni α, dobbiamo supporre che $\eta(x)$ sia continuamente derivabile e si annulli agli estremi dell'intervallo $[x_1, x_2]$.

$$\eta(x_1) = \eta(x_2) = 0. \qquad (13)$$

L'integrale (9), calcolato per $\bar{y}$, diventa una funzione del parametro α:[3]

$$I(\bar{y}) = \int_{x_1}^{x_2} F(x, y + \alpha\eta, y' + \alpha\eta')\, dx = \Phi(\alpha).$$

Poiché $y(x)$ rende minimo l'integrale, la funzione $\Phi(\alpha)$ deve avere un minimo per $\alpha = 0$; pertanto in quel punto la sua derivata deve annullarsi:

$$\Phi'(0) = \int_{x_1}^{x_2} [F_y(x, y, y')\,\eta + F_{y'}(x, y, y')\,\eta']\, dx = 0. \qquad (14)$$

Quest'ultima equazione deve essere soddisfatta per ogni funzione $\eta(x)$ continuamente derivabile e nulla agli estremi dell'intervallo $[x_1, x_2]$. Per ricavarne le conseguenze, conviene trasformare il secondo termine della condizione (14) mediante un'integrazione per parti:

$$\int_{x_1}^{x_2} F_{y'}\eta'\, dx = -\int_{x_1}^{x_2} \eta \frac{d}{dx} F_{y'}\, dx$$

così la condizione (14) assume la nuova forma

$$\Phi'(0) = \int_{x_1}^{x_2} \left(F_y - \frac{d}{dx} F_{y'}\right) \eta\, dx = 0. \qquad (15)$$

Vale il seguente semplice lemma.

Siano soddisfatte le due condizioni:

1. La funzione $f(x)$ è continua nell'intervallo $[a, b]$;

2. La funzione $\eta(x)$ è continuamente derivabile nell'intervallo $[a, b]$ e nulla ai suoi estremi.

Se, per ogni funzione $\eta(x)$ di questo tipo, l'integrale $\int_a^b f(x)\,\eta(x)\,dx$ è uguale a zero, allora $f(x) \equiv 0$.

Supponiamo infatti che in un punto c la funzione f sia diversa da zero e mostriamo che in tal caso esiste necessariamente una funzione $\eta(x)$ per la quale $\int_a^b f(x)\,\eta(x)\,dx \neq 0$, in contraddizione con l'ipotesi del lemma.

Poiché $f(c) \neq 0$ e f è continua, esiste un intorno $[\alpha, \beta]$ di c in cui f è ovunque diversa da zero e mantiene quindi un segno costante.

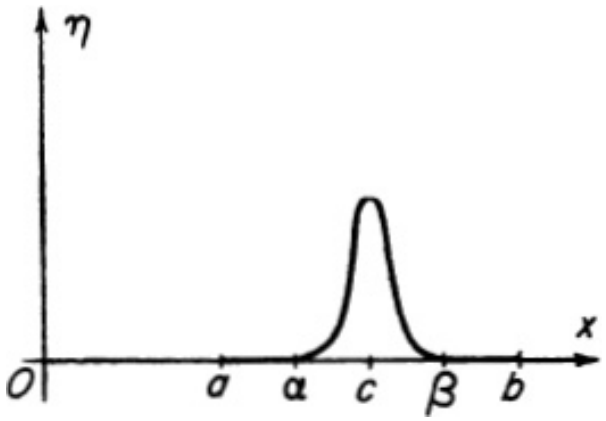


Figura 4.

Possiamo sempre costruire una funzione $\eta(x)$ continuamente derivabile in $[a, b]$, positiva nell'interno di $[\alpha, \beta]$ e nulla fuori di tale intervallo (figura 4).

Una funzione di questo tipo è, per esempio, definita dalle espressioni

$$\eta(x) = \begin{cases} 0 & \text{on } [a, \alpha], \\ (x-\alpha)^2(\beta-x)^2 & \text{on } [\alpha, \beta], \\ 0 & \text{on } [\beta, b]. \end{cases}$$

Nella formula, «on» significa «su», cioè nell'intervallo indicato.

Per una tale funzione $\eta(x)$ si ha però

$$\int_a^b f\eta \, dx = \int_a^\beta f\eta \, dx.$$

L'ultimo integrale non può essere nullo, poiché nell'interno dell'intervallo d'integrazione il prodotto $f\eta$ è diverso da zero e non cambia mai segno.

L'equazione (15) deve valere per ogni $\eta(x)$ continuamente derivabile e nulla agli estremi di $[x_1, x_2]$. In virtù del lemma, purché sussista la regolarità necessaria, ciò può accadere soltanto se

$$F_y - \frac{d}{dx} F_{y'} = 0, \qquad (16)$$

ossia, calcolando la derivata rispetto a x,

$$F_y(x, y, y') - F_{xy'}(x, y, y') - F_{yy'}(x, y, y')y' - F_{y'y'}(x, y, y')y'' = 0. \quad (17)$$

Questa è un'equazione differenziale del secondo ordine rispetto alla funzione y. Essa prende il nome di equazione di Eulero.

Possiamo formulare la conclusione seguente.

Se una funzione $y(x)$ rende minimo l'integrale $I(y)$ ed è sufficientemente regolare, deve soddisfare l'equazione differenziale di Eulero (17). Nel calcolo delle variazioni questa

affermazione ha un significato del tutto analogo alla condizione necessaria $df = 0$ nella teoria degli estremi delle funzioni. Essa consente di escludere immediatamente tutte le funzioni ammissibili che non la soddisfano, poiché per esse l'integrale non può essere minimo. L'insieme delle funzioni da studiare si restringe così considerevolmente.

Le soluzioni dell'equazione (17) hanno la proprietà che la derivata $[(d/d\alpha)I(y + \alpha\eta)]_{\alpha=0}$ è nulla per ogni $\eta(x)$; il loro significato è dunque analogo a quello dei punti stazionari di una funzione. Per questo si dice spesso che, sulle soluzioni della (17), l'integrale $I(y)$ assume un valore stazionario.

Nel nostro problema con valori assegnati agli estremi non dobbiamo trovare tutte le soluzioni dell'equazione di Eulero, ma soltanto quelle che nei punti x_1, x_2 assumono rispettivamente i valori y_1, y_2.

Richiamiamo l'attenzione sul fatto che l'equazione di Eulero (17) è del secondo ordine. Nei casi non degeneri, la sua soluzione generale contiene due costanti arbitrarie:

$$y = \phi(x, C_1 , C_2).$$

Esse devono essere determinate in modo che la curva integrale passi per i punti A e B. Otteniamo quindi, per trovare C_1 e C_2, le due equazioni

$$\phi(x_1 , C_1 , C_2) = y_1 , \quad \phi(x_2 , C_1 , C_2) = y_2 .$$

In molti casi questo sistema ha una sola soluzione, e allora esiste una sola curva integrale passante per A e B.

La ricerca delle funzioni candidate a rendere minimo l'integrale si riduce dunque al seguente problema ai limiti per equazioni differenziali: trovare nell'intervallo $[x_1, x_2]$ le soluzioni della (17) che assumono agli estremi i valori assegnati y_1, y_2.

Spesso quest'ultimo problema può essere risolto con metodi noti della teoria delle equazioni differenziali.

Ribadiamo che ogni soluzione di tale problema ai limiti è soltanto una candidata al minimo: occorre verificare se renda davvero minimo l'integrale. In casi particolari, soprattutto nelle applicazioni, l'equazione di Eulero risolve però interamente il problema. Supponiamo di sapere già che esiste una funzione che rende minimo l'integrale e che, inoltre, l'equazione di Eulero (17) abbia una sola soluzione soddisfacente le condizioni al contorno (11). Se il minimo possiede la regolarità richiesta, una sola curva ammissibile può allora essere candidata; possiamo quindi essere certi che la soluzione trovata della (17) rende effettivamente minimo l'integrale.

Esempio.

Abbiamo visto che il problema della curva di discesa più rapida si riduce a trovare il minimo dell'integrale

$$I(y) = \int_0^{x_2} \frac{\sqrt{1 + y'^2}}{\sqrt{y}}\, dx$$

nell'insieme delle funzioni che soddisfano le condizioni al contorno

$$y(0) = 0, \quad y(x_2) = y_2 .$$

In questo problema

$$F = \frac{\sqrt{1 + y'^2}}{\sqrt{y}} .$$

L'equazione di Eulero assume la forma

$$-\frac{1}{2} y^{-3/2} \sqrt{1 + y'^2} - \frac{d}{dx}\left[y^{-1/2} \frac{y'}{\sqrt{1 + y'^2}} \right] = 0.$$

che, dopo alcune trasformazioni, diventa

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

Moltiplicando entrambi i membri per y' e integrando, otteniamo

$$\ln(1 + y'^2) = -\ln y + \ln k,$$

ossia

$$y'^2 = \frac{k}{y} - 1,$$

$$\sqrt{\frac{y}{k - y}}\, dy = \pm\, dx.$$

Ponendo ora

$$y = \frac{k}{2}(1 - \cos u), \quad dy = \frac{k}{2} \sin u\, du,$$

dopo aver sostituito e semplificato, troviamo

$$\frac{k}{2}(1 - \cos u)\, du = \pm\, dx,$$

da cui, integrando, si ricava $x = \pm\, (k/2)(u - \sin u) + C$. Poiché la curva deve passare per l'origine, scegliendo $u = 0$ in quel punto si deve porre $C = 0$. Vediamo così che la brachistocrona è la cicloide

$$x = \frac{k}{2}(u - \sin u), \quad y = \frac{k}{2}(1 - \cos u).$$

La costante k si determina imponendo che la curva passi per $M_2(x_2, y_2)$.

Funzionali dipendenti da più funzioni.

Il funzionale più semplice del calcolo delle variazioni, dato dalla (9), dipendeva da una sola funzione. Nelle applicazioni tali funzionali compaiono quando gli oggetti, o il loro comportamento, sono descritti da un'unica dipendenza funzionale. Per esempio, una curva piana è descritta dalla dipendenza dell'ordinata di un punto dalla sua ascissa; il moto di un punto materiale lungo un asse è descritto dalla dipendenza della sua coordinata dal tempo, e così via.

Spesso, però, dobbiamo considerare oggetti che non possono essere descritti così semplicemente. Per definire una curva nello spazio occorre conoscere, quando questa rappresentazione è possibile, la dipendenza di due coordinate dalla terza. Il moto di un punto nello spazio è definito dalla dipendenza delle sue tre coordinate dal tempo, e così via. Lo studio di questi oggetti più complessi conduce a problemi variazionali in cui variano più funzioni.

Ci limiteremo al caso in cui il funzionale dipende da due funzioni $y(x)$ e $z(x)$, poiché il caso di un numero maggiore di funzioni non presenta differenze di principio.

Consideriamo il problema seguente. Le coppie ammissibili di funzioni $y(x)$ e $z(x)$ sono definite dalle condizioni:

1. Le funzioni

$$y = y(x), \quad z = z(x) \tag{18}$$

sono continuamente derivabili nell'intervallo $[x_1, x_2]$;

2. Agli estremi dell'intervallo esse assumono valori assegnati:

$$\begin{aligned} y(x_1) &= y_1\,, \quad y(x_2) = y_2\,, \\ z(x_1) &= z_1\,, \quad z(x_2) = z_2\,. \end{aligned} \tag{19}$$

Fra tutte le coppie di funzioni $y(x)$ e $z(x)$ ammissibili, dobbiamo trovare quella che rende minimo l'integrale

$$I(y, z) = \int_{x_1}^{x_2} F(x, y, z, y', z')\, dx. \tag{20}$$

Nello spazio tridimensionale x, y, z, a ogni coppia ammissibile corrisponde una curva l, definita dalle equazioni (18) e passante per i punti

$$M_1(x_1\,, y_1\,, z_1), \quad M_2(x_2\,, y_2\,, z_2).$$

Dobbiamo trovare il minimo dell'integrale (20) nell'insieme di tutte queste curve.

Supponiamo che esista una coppia di funzioni che rende minimo l'integrale (20) e indichiamola con $y(x)$, $z(x)$. Accanto a essa consideriamo una seconda coppia di funzioni,

$$\bar{y} = y + \alpha\eta(x), \quad \bar{z} = z + \alpha\zeta(x),$$

dove $\eta(x)$ e $\zeta(x)$ sono funzioni arbitrarie continuamente derivabili e nulle agli estremi x_1, x_2 dell'intervallo. Anche $\bar{y}$, $\bar{z}$ sono ammissibili e, per $\alpha = 0$, coincidono con y, z. Sostituiamole nella (20):

$$I(\bar{y}, \bar{z}) = \int_{x_1}^{x_2} F(x, y + \alpha\eta, z + \alpha\zeta, y' + \alpha\eta', z' + \alpha\zeta')\, dx = \Phi(\alpha).$$

L'integrale ottenuto è una funzione di α. Poiché $\bar{y}$ e $\bar{z}$ coincidono con y e z per $\alpha = 0$, la funzione $\Phi(\alpha)$ deve avere un minimo per $\alpha = 0$. In un punto di minimo la derivata di Φ deve però annullarsi:

$$\Phi'(0) = 0.$$

Calcolando la derivata, si ottiene

$$\int_{x_1}^{x_2} (F_y \cdot \eta + F_z \cdot \zeta + F_{y'} \cdot \eta' + F_{z'} \cdot \zeta')\, dx = 0,$$

oppure, integrando per parti i termini contenenti η' e ζ',

$$\int_{x_1}^{x_2} \left[\left(F_y - \frac{d}{dx} F_{y'}\right) \eta(x) + \left(F_z - \frac{d}{dx} F_{z'}\right) \zeta(x)\right] dx = 0.$$

Quest'ultima equazione deve essere soddisfatta per qualsiasi coppia di funzioni $\eta(x)$, $\zeta(x)$ continuamente derivabili e nulle agli estremi dell'intervallo. In virtù del lemma fondamentale dimostrato in precedenza, devono dunque valere le due condizioni

$$\begin{aligned} F_y - \frac{d}{dx} F_{y'} &= 0, \\ F_z - \frac{d}{dx} F_{z'} &= 0. \end{aligned} \qquad (21)$$

Pertanto, se le funzioni y, z rendono minimo l'integrale (20), nelle ipotesi di regolarità considerate esse devono soddisfare il sistema di equazioni differenziali di Eulero (21).

Anche questo risultato consente di ricondurre la ricerca delle candidate al minimo dell'integrale (20) a un problema ai limiti della teoria delle equazioni differenziali: nell'intervallo $[x_1, x_2]$ dobbiamo trovare le soluzioni y, z del sistema (21) che soddisfano le condizioni al contorno (19).

Come nel caso precedente, si apre così una possibile via per risolvere il problema di minimo.

Come esempio di applicazione del sistema di Eulero (21), consideriamo il principio variazionale di Ostrogradskij-Hamilton nella meccanica newtoniana, limitandoci alla sua forma più semplice.

Consideriamo un corpo materiale di massa m e supponiamo di poterne trascurare dimensioni e forma, così da trattarlo come un punto materiale.

Supponiamo che il punto si muova dalla posizione $M_1(x_1, y_1, z_1)$ all'istante t_1 alla posizione $M_2(x_2, y_2, z_2)$ all'istante t_2. Supponiamo inoltre che il moto obbedisca alle leggi della meccanica newtoniana e sia prodotto da una forza $F(x, y, z, t)$, dipendente dalla posizione e dal tempo e dotata di una funzione potenziale $U(x, y, z, t)$. Quest'ultima ipotesi significa che le componenti F_x, F_y, F_z della forza lungo gli assi sono le derivate parziali di U rispetto alle coordinate corrispondenti:

$$F_x = \frac{\partial U}{\partial x}, \quad F_y = \frac{\partial U}{\partial y}, \quad F_z = \frac{\partial U}{\partial z}.$$

Supponiamo che il moto sia libero, cioè non soggetto ad alcun vincolo.[4]

Le equazioni newtoniane del moto sono

$$m\frac{d^2x}{dt^2} = \frac{\partial U}{\partial x}, \quad m\frac{d^2y}{dt^2} = \frac{\partial U}{\partial y}, \quad m\frac{d^2z}{dt^2} = \frac{\partial U}{\partial z}.$$

Se il punto obbedisce alle leggi della meccanica newtoniana, il suo moto è determinato dalle appropriate condizioni iniziali. Accanto a questi «moti newtoniani», consideriamo però altri moti, non necessariamente newtoniani, che per brevità chiameremo «ammissibili». Essi sono soggetti soltanto alle due condizioni di trovarsi nella posizione M_1 all'istante t_1 e nella posizione M_2 all'istante t_2, oltre alla regolarità necessaria.

Come possiamo distinguere il «moto newtoniano» del punto dagli altri moti «ammissibili»? Una risposta è data dal principio di Ostrogradskij-Hamilton.

Introduciamo l'energia cinetica del punto,

$$T = \tfrac{1}{2}m(x'^2 + y'^2 + z'^2)$$

e costruiamo il cosiddetto integrale d'azione,

$$I = \int_{t_1}^{t_2} (T + U)\, dt.$$

Il principio afferma: il «moto newtoniano» del punto si distingue fra tutti i moti «ammissibili» perché rende stazionario l'integrale d'azione.

L'integrale d'azione I dipende da tre funzioni: $x(t)$, $y(t)$, $z(t)$.

Poiché per tutti i moti messi a confronto le posizioni iniziale e finale sono le stesse, i valori agli estremi di queste funzioni sono fissati. Si tratta di un problema variazionale per tre funzioni variabili, con valori assegnati agli estremi dell'intervallo $[t_1, t_2]$.

Abbiamo convenuto di dire che l'integrale (9) assume un valore stazionario su ogni curva integrale dell'equazione di Eulero. Nel nostro problema la funzione integranda è

$$F = T + U = \tfrac{1}{2}m(x'^2 + y'^2 + z'^2) + U(x, y, z, t)$$

e dipende da tre funzioni incognite. Perché l'integrale sia stazionario deve quindi essere soddisfatto il sistema di tre equazioni differenziali

$$F_x - \frac{d}{dt}F_{x'} = 0,$$
$$F_y - \frac{d}{dt}F_{y'} = 0,$$
$$F_z - \frac{d}{dt}F_{z'} = 0.$$

Per questa funzione integranda si ha $F_x = \partial U/\partial x$ e $F_{x'} = mx'$, con espressioni analoghe per y e z. Il sistema di Eulero coincide dunque con le equazioni del moto della meccanica newtoniana: ciò verifica il principio di Ostrogradskij-Hamilton.

Il problema del minimo di un integrale multiplo.

L'ultimo problema del calcolo delle variazioni sul quale vogliamo richiamare l'attenzione del lettore è la minimizzazione di un integrale multiplo. Poiché i fatti connessi alla soluzione sono analoghi per integrali di qualunque molteplicità, ci limiteremo al caso più semplice, quello degli integrali doppi.

Sia B un dominio del piano Oxy, delimitato dal contorno l. L'insieme delle funzioni ammissibili è definito dalle condizioni:

1. $u(x, y)$ è continuamente derivabile nel dominio B;

2. Sul contorno l la funzione u assume valori assegnati:

$$u\,|_l = f(M). \qquad (22)$$

Fra tutte queste funzioni dobbiamo trovare quella che rende minimo l'integrale

$$I(u) = \iint_B F(x, y, u, u_x, u_y)\,dx\,dy. \qquad (23)$$

Nello spazio (x, y, u), i valori al contorno (22) assegnati a u determinano una curva spaziale Γ situata sopra l (cfr. figura 2 di questo capitolo). Consideriamo tutte le superfici S che sono grafici sopra B e passano per Γ. Fra esse vogliamo trovare quella per cui l'integrale (23) è minimo.

Come in precedenza, supponiamo che la funzione minimizzante esista e indichiamola con u. Accanto a essa consideriamo un'altra funzione,

$$\bar{u} = u + \alpha\eta(x, y),$$

dove $\eta(x, y)$ è una funzione arbitraria continuamente derivabile e nulla su l. Allora la funzione

$$I(\bar{u}) = \iint_B F(x, y, u + \alpha\eta, u_x + \alpha\eta_x, u_y + \alpha\eta_y)\, dx\, dy = \Phi(\alpha)$$

deve avere un minimo per $\alpha = 0$. In tal caso, per $\alpha = 0$, la sua derivata prima deve essere nulla:

$$\Phi'(0) = 0,$$

ossia

$$\iint_B (F_u\eta + F_{u_x}\eta_x + F_{u_y}\eta_y)\, dx\, dy = 0. \qquad (24)$$

Trasformiamo gli ultimi due termini mediante la formula di Ostrogradskij:

$$\begin{aligned}
&\iint_B (F_{u_x}\eta_x + F_{u_y}\eta_y)\, dx\, dy \\
&= \iint_B \left[\frac{\partial}{\partial x}(F_{u_x}\eta) + \frac{\partial}{\partial y}(F_{u_y}\eta)\right] dx\, dy - \iint_B \left(\frac{\partial}{\partial x}F_{u_x} + \frac{\partial}{\partial y}F_{u_y}\right)\eta\, dx\, dy \\
&= \int_l [F_{u_x}\cos(n, x) + F_{u_y}\cos(n, y)]\,\eta\, ds \\
&\qquad - \iint_B \left(\frac{\partial}{\partial x}F_{u_x} + \frac{\partial}{\partial y}F_{u_y}\right)\eta\, dx\, dy.
\end{aligned}$$

L'integrale curvilineo lungo l si annulla, poiché sul contorno l la funzione η è nulla. La condizione (24) si può quindi scrivere nella forma

$$\iint_B \left(F_u - \frac{\partial}{\partial x}F_{u_x} - \frac{\partial}{\partial y}F_{u_y}\right)\eta\, dx\, dy = 0.$$

Questa equazione deve essere soddisfatta per ogni funzione η continuamente derivabile e nulla sul contorno l.

Come prima, possiamo concludere che in tutti i punti del dominio B deve valere l'equazione

$$F_u - \frac{\partial}{\partial x}F_{u_x} - \frac{\partial}{\partial y}F_{u_y} = 0 \qquad (25)$$

nelle ipotesi di regolarità necessarie.

Dunque, se la funzione u rende minimo l'integrale (23), deve soddisfare l'equazione alle derivate parziali (25).

Come in tutti i problemi precedenti, abbiamo stabilito un legame fra il problema variazionale di minimizzare un integrale e un problema ai limiti per un'equazione differenziale, in questo caso alle derivate parziali.

Esempio.

Lo spostamento $u(x, y)$ dei punti di una membrana il cui bordo è deformato si determina imponendo il minimo dell'energia potenziale

$$\frac{\mu}{2} \iint_B (u_x^2 + u_y^2)\, dx\, dy$$

per i valori al contorno assegnati $u|_l = \phi$.

Omettendo per semplicità il fattore costante μ, possiamo porre

$$F = \frac{1}{2}(u_x^2 + u_y^2),$$

e l'equazione (25) assume la forma

$$-\frac{\partial}{\partial x} u_x - \frac{\partial}{\partial y} u_y = 0,$$

ossia

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Il problema di determinare lo spostamento dei punti della membrana è così ricondotto a quello di trovare una funzione armonica u con valori assegnati sul contorno del dominio (cfr. capitolo VI, §3).

§3. Metodi di soluzione approssimata dei problemi del calcolo delle variazioni

Concludiamo il capitolo illustrando le idee alla base di alcuni metodi di approssimazione del calcolo delle variazioni.

Per fissare le idee, consideriamo il funzionale più semplice,

$$I(y) = \int_{x_1}^{x_2} F(x, y, y')\, dx$$

con valori assegnati agli estremi per le funzioni ammissibili.

Sia $y(x)$ una soluzione esatta del problema di minimizzazione di I, e sia $m = I(y)$ il corrispondente valore minimo dell'integrale. Sembrerebbe che, trovando una funzione ammissibile $\bar{y}$ per la quale il valore dell'integrale $I(\bar{y})$ sia molto vicino a m, si possa supporre che anche $\bar{y}$ differisca poco dalla soluzione esatta y. Inoltre, se riusciamo a costruire una successione di funzioni ammissibili $\bar{y}_1, \bar{y}_2, \ldots$ per le quali $I(\bar{y}_n) \to m$, possiamo sperare che tale successione converga, in un senso opportuno, alla soluzione y. Il calcolo di $\bar{y}_n$ con indice abbastanza grande permetterebbe allora di trovare la soluzione con qualunque precisione desiderata.

A seconda di come si sceglie la «successione minimizzante» $\bar{y}_n (n = 1, 2, \cdots)$, si ottengono i diversi metodi di approssimazione del calcolo delle variazioni.

Storicamente, il primo fu il metodo delle spezzate, o metodo di Eulero. Suddividiamo l'intervallo $[x_1, x_2]$ in un certo numero di segmenti. Se, per esempio, li scegliamo di uguale lunghezza, i punti di suddivisione sono

$$x_1 , x_1 + h, x_1 + 2h, \cdots, x_1 + nh = x_2 , h = \frac{x_2 - x_1}{n} .$$

Costruiamo ora la spezzata p_{n-1} con i vertici situati sopra i punti di suddivisione. Indichiamo le ordinate dei vertici con

$$b_0 , b_1 , b_2 , \cdots , b_{n-1} , b_n$$

e imponiamo che la spezzata inizi e termini negli stessi punti delle curve ammissibili, cosicché $b_0 = y_1$ e $b_n = y_2$. La spezzata è allora determinata dalle ordinate

$$b_1 , b_2 , \cdots , b_{n-1} .$$

Si tratta ora di scegliere la spezzata p_{n-1}, ossia le ordinate b_i dei suoi vertici, in modo da approssimare quanto meglio possibile la soluzione esatta del problema.

A questo scopo è naturale procedere così. Calcoliamo l'integrale I sulla spezzata. Il suo valore dipende dai b_i:

$$I(p_{n-1}) = \Phi(b_1, b_2, \cdots, b_{n-1})$$

È dunque una funzione di queste ordinate. Scegliamo ora i b_i in modo da rendere minimo $I(p_{n-1})$. Per determinarli otteniamo il sistema di condizioni necessarie

$$\frac{\partial}{\partial b_i} I(p_{n-1}) = 0 \quad (i = 1, 2, \cdots, n-1).$$

Ogni curva ammissibile, e in particolare la soluzione esatta, può essere approssimata da spezzate con precisione arbitraria sia nella posizione nel piano sia nelle direzioni delle tangenti, considerate sui singoli segmenti. Sotto le opportune ipotesi sull'integrando, la successione di spezzate p_{n-1} così costruita è pertanto una successione minimizzante. Prendendo n abbastanza grande, possiamo sperare di approssimare la soluzione con qualsiasi precisione desiderata sull'intero intervallo $[x_1, x_2]$. Naturalmente, la convergenza deve essere studiata in ogni singolo caso.

Il metodo seguente, molto comodo per il calcolo, è ampiamente usato nella fisica e nella tecnica.

Scegliamo una funzione $\phi_0(x)$ che soddisfi le condizioni al contorno $\phi_0(x_1) = y_1$ e $\phi_0(x_2) = y_2$, e una successione di funzioni $\phi_1(x)$, $\phi_2(x)$, ... nulle agli estremi di $[x_1, x_2]$.

Formiamo quindi la combinazione lineare

$$s_n(x) = \phi_0(x) + a_1\phi_1(x) + \cdots + a_n\phi_n(x).$$

Per qualunque valore dei coefficienti numerici $a_1, a_2, \ldots, a_n$, la funzione $s_n(x)$ è ammissibile.

Sostituendo $s_n(x)$ a y nell'integrale I ed eseguendo i calcoli, otteniamo una funzione dei coefficienti a_i.

Scegliamo ora questi coefficienti in modo che la funzione assuma il valore minimo possibile.

I coefficienti devono soddisfare il sistema

$$\frac{\partial}{\partial a_i} I(s_n) = 0 \quad (i = 1, 2, \cdots, n).$$

Risolvendo il sistema e individuando fra le soluzioni quelle di minimo, otteniamo i valori dei coefficienti $a_1, \ldots, a_n$ che rendono minimo $I(s_n)$. Con essi costruiamo l'approssimazione della soluzione

$$\bar{s}_n(x) = \phi_0(x) + \bar{a}_1\phi_1(x) + \cdots + \bar{a}_n\phi_n(x).$$

La successione di approssimazioni $\bar{s}_n$ $(n = 1, 2, \cdots)$ così costruita non è minimizzante per una scelta arbitraria delle funzioni ϕ. Perché il metodo consenta l'approssimazione generale

occorre che la successione delle funzioni ϕ_i soddisfi un'opportuna condizione di «completezza», che qui non preciseremo.

Letture consigliate

G. A. Bliss, Calculus of variations, Open Court, La Salle, Illinois, 1925.

C. Lanczos, The variational principles of mechanics, University of Toronto Press, 1949.

G. Pólya e G. Szegö, Isoperimetric inequalities in mathematical physics, Princeton University Press, 1951.

R. Weinstock, Calculus of variations, with applications to physics and engineering, McGraw-Hill, New York, 1952.

Note al capitolo VIII

[1] L'origine del nome «calcolo delle variazioni» viene spiegata più avanti.

[2] Qui e in tutto il capitolo usiamo indici per indicare gli argomenti rispetto ai quali si calcolano le derivate parziali.

[3] La differenza $\bar{y} - y = \alpha\eta$ si chiama variazione, ossia cambiamento, della funzione y e si indica con δy; la differenza $I(\bar{y}) - I(y)$ si chiama variazione totale dell'integrale (9). Da qui deriva il nome «calcolo delle variazioni».

[4] Questa ipotesi non è essenziale per il principio di Ostrogradskij-Hamilton: si possono imporre al sistema meccanico vincoli anche dipendenti dal tempo, purché olonomi, cioè descrivibili mediante equazioni che non contengano derivate temporali delle coordinate. Si intendono vincoli ideali e variazioni compatibili con essi.

Nota editoriale

La traduzione conserva le tre sezioni, le quattro figure, le formule e le quattro note del capitolo. Sono corretti alcuni refusi nelle coordinate dei punti, nei rinvii alle formule e nella bibliografia. Nel calcolo della cicloide è ripristinato il segno meno fra u e sin u, coerentemente con la formula finale. Nella verifica del principio di Hamilton la derivata dell'integranda rispetto a x′ è mx′. Il richiamo alla nota 3, non leggibile accanto alla formula originale, è riportato nella frase che la precede.

La brachistocrona è considerata in un campo gravitazionale uniforme, senza attrito. Nell'interno dell'arco si richiede y > 0; l'integrale del tempo all'origine è improprio. La cicloide ha tangente verticale nel punto iniziale: la derivabilità continua come grafico non va quindi richiesta anche a quell'estremo. Il calcolo con l'equazione di Eulero individua la curva candidata; la verifica del minimo e la scelta dell'arco ammissibile richiedono argomenti ulteriori. Nella formula dell'area di rotazione si suppone y ≥ 0; in generale occorre il valore assoluto di y.

Le integrazioni per parti e le forme classiche delle equazioni di Eulero richiedono regolarità sufficiente delle funzioni e dei contorni. Il coefficiente della derivata seconda può annullarsi: l'affermazione sulla soluzione generale con due costanti riguarda il caso non degenere. Le equazioni di Eulero esprimono condizioni necessarie di stazionarietà e non garantiscono da sole un minimo. Nel principio di Ostrogradskij-Hamilton si conserva la convenzione dell'originale: la forza è il gradiente di U, quindi U è l'opposto dell'energia potenziale usuale e l'integranda T + U corrisponde a T − V.

Nel §3 la vicinanza dei valori del funzionale al minimo non implica, da sola, la vicinanza delle funzioni a una soluzione. La convergenza richiede ipotesi ulteriori e una precisa scelta del senso di convergenza. Nel metodo delle spezzate si ammettono funzioni derivabili a tratti; i vertici non impediscono di calcolare l'integrale. I sistemi ottenuti annullando le derivate rispetto ai coefficienti danno condizioni necessarie: fra le loro soluzioni occorre ancora individuare i minimi. La completezza delle funzioni di prova riguarda l'approssimazione nello spazio funzionale appropriato.

CAPITOLO IX

Funzioni di variabile complessa

M. V. Keldyš

§1. Numeri complessi e funzioni di variabile complessa

I numeri complessi e il loro significato in algebra.

I numeri complessi furono introdotti nella matematica in relazione alla soluzione delle equazioni algebriche. L'impossibilità di risolvere l'equazione

$$x^2 + 1 = 0 \qquad (1)$$

nel campo dei numeri reali condusse a introdurre convenzionalmente un nuovo numero, l'unità immaginaria i, definita dalla relazione

$$i^2 = -1. \qquad (2)$$

I numeri della forma $a + bi$, con a e b reali, furono chiamati numeri complessi. Si operava su di essi come sui numeri reali, sommandoli e moltiplicandoli come binomi. Usando anche la relazione (2), le operazioni aritmetiche fondamentali applicate a numeri complessi producono ancora numeri complessi.[1] La divisione, definita come operazione inversa della moltiplicazione, risulta anch'essa univocamente determinata, purché il denominatore non sia nullo. L'introduzione dei numeri complessi mise così in luce un fatto interessante, benché inizialmente soltanto formale: oltre ai numeri reali esistono altri numeri sui quali si possono eseguire tutte le operazioni aritmetiche.

Il passo successivo fu la rappresentazione geometrica. Ogni numero complesso $a + bi$ può essere rappresentato da un punto del piano Oxy di coordinate (a, b), oppure da un vettore che congiunge l'origine a quel punto. Ne derivò un nuovo punto di vista: i numeri complessi sono coppie (a, b) di numeri reali per le quali si definiscono addizione e moltiplicazione in modo da rispettare le stesse leggi valide per i reali. Qui emerge un fatto notevole. La somma di due numeri complessi

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

è rappresentata geometricamente dalla diagonale del parallelogramma costruito sui vettori che rappresentano gli addendi (figura 1). I numeri complessi si sommano dunque secondo la stessa legge delle grandezze vettoriali della meccanica e della fisica: forze, velocità e accelerazioni. Fu un ulteriore motivo per considerarli non soltanto generalizzazioni formali, ma strumenti capaci di rappresentare grandezze fisiche reali.

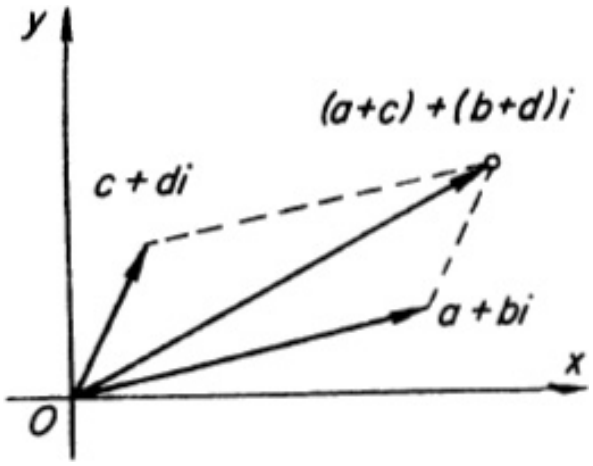


Figura 1.

Vedremo in seguito quanto sia fecondo questo punto di vista in vari problemi della fisica matematica.

I primi successi dei numeri complessi riguardarono però la scoperta delle leggi dell'algebra e dell'analisi. Il campo dei numeri reali, chiuso rispetto alle operazioni aritmetiche, si rivelò insufficiente per l'algebra. Persino un'equazione semplice come la (1) non ha radici reali; nel campo complesso vale invece il fatto notevole espresso dal teorema fondamentale dell'algebra: ogni equazione algebrica

$$z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0$$

di grado n, con coefficienti complessi, ha n radici complesse, contate con la loro molteplicità.[2]

Il teorema mostra che i numeri complessi formano un sistema, in un senso ben preciso, completo rispetto alle operazioni dell'algebra. Non è affatto ovvio che aggiungere ai reali una radice della sola equazione (1) conduca ai numeri $a + bi$, nel cui campo ogni equazione algebrica è risolubile. Il teorema fondamentale dell'algebra mostrò che la teoria dei polinomi, anche a coefficienti reali, acquista una forma compiuta soltanto considerando i valori del polinomio nell'intero piano complesso. Gli sviluppi successivi confermarono sempre più questo punto di vista: le proprietà dei polinomi si rivelano pienamente quando li si considera come funzioni di variabile complessa.

Serie di potenze e funzioni di variabile complessa.

Lo sviluppo dell'analisi portò alla luce una serie di fatti che mostravano l'importanza dei numeri complessi non soltanto per i polinomi, ma anche per un'altra classe fondamentale di funzioni: quelle sviluppabili in serie di potenze,

$$f(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots. \qquad (3)$$

Come si è ricordato nel capitolo II, lo sviluppo dell'analisi infinitesimale richiese una definizione più precisa del concetto di funzione e dei diversi modi di assegnarla. Senza soffermarci su questi interessanti problemi, ricordiamo che fin dagli inizi dell'analisi risultò

possibile sviluppare in serie di potenze, nell'intorno dei punti regolari del loro dominio, le funzioni incontrate più frequentemente. Questa proprietà vale, per esempio, per le funzioni elementari nei punti in cui sono analitiche.

La maggior parte dei problemi concreti dell'analisi conduceva a funzioni sviluppabili in serie di potenze. D'altra parte, si desiderava collegare la definizione di funzione «matematica» a una formula «matematica», e le serie di potenze costituivano una forma molto generale di tali formule. Si tentò quindi seriamente di limitare l'analisi allo studio delle funzioni sviluppabili in serie di potenze, dette funzioni analitiche. Lo sviluppo della scienza mostrò l'inopportunità di questa restrizione. I problemi della fisica matematica cominciarono a oltrepassare la classe delle funzioni analitiche, che non comprende neppure, per esempio, le funzioni rappresentate da curve con uno spigolo. Tuttavia, per le sue proprietà notevoli e le numerose applicazioni, la classe delle funzioni analitiche rimase una delle più importanti studiate dai matematici.

Poiché il calcolo di ciascun termine di una serie di potenze richiede soltanto operazioni aritmetiche, una funzione rappresentata da una tale serie può essere calcolata anche per valori complessi dell'argomento, almeno là dove la serie converge. Estendendo così una funzione reale ad argomenti complessi, si parla di «prolungamento» della funzione nel campo complesso. Una funzione analitica, come un polinomio, può quindi essere considerata per argomenti reali e complessi. Si possono inoltre considerare serie di potenze con coefficienti complessi. Le proprietà delle funzioni analitiche, come quelle dei polinomi, si manifestano pienamente soltanto nel campo complesso. Vediamolo con un esempio.

Consideriamo le due funzioni di variabile reale

$$e^x \text{ e } 1/(1+x^2).$$

Entrambe sono finite, continue e derivabili un numero arbitrario di volte su tutto l'asse Ox. Possiamo svilupparle in serie di Taylor, per esempio intorno all'origine $x = 0$:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots, \qquad (4)$$

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots. \qquad (5)$$

La prima serie converge per ogni x, mentre la seconda converge soltanto per $-1 < x < 1$. Considerando la funzione della (5) per argomenti reali, non si comprende perché la sua serie di Taylor diverga per $|x| \geqslant 1$. Il passaggio al campo complesso chiarisce la situazione. Consideriamo la serie (5) per argomenti complessi:

$$1 - z^2 + z^4 - z^6 + \cdots. \qquad (6)$$

La somma di n termini di questa serie,

$$s_n = 1 - z^2 + z^4 - z^6 + \cdots + (-1)^{n-1} z^{2n-2}$$

si calcola come nel caso reale:

$$s_n + z^2 s_n = 1 - (-1)^n z^{2n},$$

pertanto

$$s_n = [1 - (-1)^n z^{2n}]/(1 + z^2),$$

Questa espressione mostra che, per $|z| < 1$,

$$\lim_{n\to\infty} s_n = \frac{1}{1+z^2},$$

poiché $|z|^{2n} \to 0$. Per i valori complessi z con $|z| < 1$, la serie (6) converge dunque alla somma $1/(1 + z^2)$. Per $|z| \geqslant 1$ la serie diverge, perché la differenza $s_n - s_{n-1} = (-1)^{n-1} z^{2n-2}$ non tende a zero.

La disuguaglianza $|z| < 1$ significa che la distanza del punto z dall'origine è minore di uno. I punti di convergenza della (6) formano quindi un disco aperto del piano complesso con centro nell'origine. Sulla circonferenza si trovano i due punti i e $-i$, nei quali $1/(1 + z^2)$ ha un polo: la presenza di queste singolarità determina il limite del dominio di convergenza della serie.

Il dominio di convergenza di una serie di potenze.

Per una serie di potenze

$$a_0 + a_1(z - a) + a_2(z - a)^2 + \cdots + a_n(z - a)^n + \cdots \qquad (7)$$

nel piano complesso, l'interno dell'insieme di convergenza è un disco con centro nel punto a, con i possibili casi degeneri descritti sotto. Sulla circonferenza il comportamento va esaminato separatamente.

Dimostriamo questa proposizione, chiamata teorema di Abel.

Osserviamo innanzitutto che una serie con termini complessi w_n,

$$w_1 + w_2 + \cdots + w_n + \cdots, \qquad (8)$$

può essere considerata come due serie, formate rispettivamente dalle parti reali e dai coefficienti delle parti immaginarie di $w_n = u_n + iv_n$:

$$u_1 + u_2 + \cdots + u_n + \cdots, \qquad (9)$$

$$v_1 + v_2 + \cdots + v_n + \cdots. \qquad (10)$$

La somma parziale s_n della (8) si esprime mediante le somme parziali σ_n e τ_n delle (9) e (10):

$$s_n = \sigma_n + i\tau_n,$$

Quindi la convergenza della (8) equivale a quella di entrambe le serie (9) e (10), e la sua somma s si esprime mediante le loro somme σ e τ:

$$s = \sigma + i\tau.$$

Dopo queste osservazioni, è immediato il lemma seguente:

Se i moduli dei termini della serie (8) sono minori dei termini di una serie geometrica convergente

$$A + Aq + \cdots + Aq^n + \cdots$$

con $A > 0$ e $0 < q < 1$, allora la serie (8) converge.

Infatti, se $|w_n| < Aq^n$, si ha

$$|u_n| \leqslant |w_n| < Aq^n,$$
$$|v_n| \leqslant |w_n| < Aq^n,$$

e quindi, come si è visto nel capitolo II, §14, convergono le serie (9) e (10), e di conseguenza anche la (8).

Mostriamo ora che, se la serie di potenze (7) converge in un punto z_0, essa converge in tutti i punti interni alla circonferenza di centro a passante per z_0 (figura 2). Ne segue facilmente che la serie (7),

$$a_0 + a_1(z - a) + \cdots + a_n(z - a)^n + \cdots$$

converge nell'intero piano, oppure soltanto nel punto $z = a$, oppure all'interno di un disco di raggio finito; in quest'ultimo caso può convergere anche in alcuni o in tutti i punti della circonferenza.

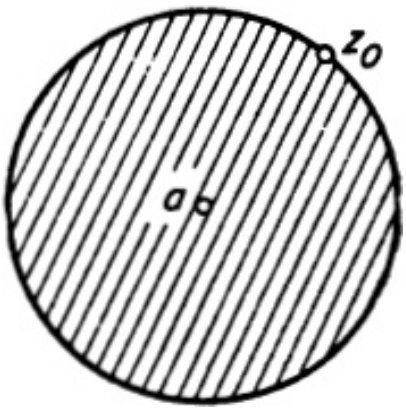


Figura 2.

Supponiamo infatti che la (7) converga in z_0. Il suo termine generale, calcolato per $z = z_0$, tende a zero per $n \to \infty$. Tutti i termini sono quindi contenuti in un disco: indicandone con A il raggio, per ogni n si ha

$$| a_n(z_0 - a)^n | < A. \qquad (11)$$

Prendiamo un punto z più vicino ad a di quanto lo sia z_0 e mostriamo che la serie converge anche lì.

Evidentemente

$$| z - a | < | z_0 - a |,$$

perciò

$$q = \frac{| z - a |}{| z_0 - a |} < 1. \qquad (12)$$

Stimiamo il termine generale della (7) nel punto z:

$$| a_n(z - a)^n | = \left| a_n(z_0 - a)^n \left(\frac{z - a}{z_0 - a}\right)^n \right| = | a_n(z_0 - a)^n | \left(\frac{| z - a |}{| z_0 - a |}\right)^n;$$

dalle disuguaglianze (11) e (12) segue che

$$| a_n(z - a)^n | < Aq^n;$$

Il modulo del termine generale della (7) è dunque minore del termine corrispondente di una serie geometrica convergente. Per il lemma precedente, la serie (7) converge nel punto z.

Il disco all'interno del quale una serie di potenze converge e all'esterno del quale diverge si chiama disco di convergenza; il suo raggio è il raggio di convergenza. Si può dimostrare che, se il raggio è finito e positivo, la circonferenza passa per almeno una singolarità della funzione più vicina al centro a, cioè per un punto oltre il quale viene meno il suo comportamento analitico regolare.

La serie (4) converge in tutto il piano complesso; la (5), come abbiamo mostrato, ha raggio di convergenza uguale a uno.

Funzioni esponenziali e trigonometriche di variabile complessa.

Una serie di potenze può servire a «prolungare» nel campo complesso una funzione reale. Per esempio, per z complesso definiamo e^z mediante la serie

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots. \qquad (13)$$

Analogamente si introducono le funzioni trigonometriche di variabile complessa:

$$\sin z = \frac{z}{1!} - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots, \qquad (14)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots . \qquad (15)$$

Queste serie convergono in tutto il piano.

È interessante osservare il legame che si manifesta fra funzioni esponenziali e trigonometriche passando al campo complesso.

Sostituendo *iz* a *z* nella (13), otteniamo

$$e^{iz} = 1 + i\frac{z}{1!} - \frac{z^2}{2!} - i\frac{z^3}{3!} + \frac{z^4}{4!} + \cdots .$$

Raggruppando i termini privi del fattore *i* e quelli che lo contengono, si ha

$$e^{iz} = \cos z + i \sin z. \qquad (16)$$

Analogamente si ricava

$$e^{-iz} = \cos z - i \sin z. \qquad (16')$$

Le relazioni (16) e (16′) si chiamano formule di Eulero. Risolvendole rispetto a cos *z* e sin *z*, otteniamo

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}. \qquad (17)$$

È molto importante che, anche per argomenti complessi, valga la semplice regola di addizione degli esponenti:

$$e^{z_1} \cdot e^{z_2} = e^{z_1+z_2} . \qquad (18)$$

Poiché per argomenti complessi abbiamo definito e^z mediante la serie (13), dobbiamo dimostrare la (18) partendo da questa definizione. Ecco la dimostrazione:

$$e^{z_1} \cdot e^{z_2} = \left(1 + \frac{z_1}{1!} + \frac{z_1^2}{2!} + \cdots\right) \cdot \left(1 + \frac{z_2}{1!} + \frac{z_2^2}{2!} + \cdots\right).$$

Eseguiamo la moltiplicazione delle serie termine per termine. I prodotti ottenuti possono essere disposti in una tabella quadrata:

$$1\cdot 1+1\cdot\frac{z_2}{1!}+1\cdot\frac{z_2^2}{2!}+1\cdot\frac{z_2^3}{3!}+$$
$$\cdots+\frac{z_1}{1!}\cdot 1+\frac{z_1}{1!}\cdot\frac{z_2}{1!}+\frac{z_1}{1!}\cdot\frac{z_2^2}{2!}+\frac{z_1}{1!}\cdot\frac{z_2^3}{3!}+$$
$$\cdots+\frac{z_1^2}{2!}\cdot 1+\frac{z_1^2}{2!}\cdot\frac{z_2}{1!}+\frac{z_1^2}{2!}\cdot\frac{z_2^2}{2!}+\frac{z_1^2}{2!}\cdot\frac{z_2^3}{3!}+$$
$$\cdots+\frac{z_1^3}{3!}\cdot 1+\frac{z_1^3}{3!}\cdot\frac{z_2}{1!}+\frac{z_1^3}{3!}\cdot\frac{z_2^2}{2!}+\frac{z_1^3}{3!}\cdot\frac{z_2^3}{3!}+\cdots.$$

Raggruppiamo ora i termini per i quali la somma degli esponenti di z_1 e z_2 è la stessa. Come si vede, essi si trovano sulle diagonali della tabella. Otteniamo

$$e^{z_1}\cdot e^{z_2}=1+\left(\frac{z_2}{1!}+\frac{z_1}{1!}\right)+\left(\frac{z_2^2}{2!}+\frac{z_1}{1!}\frac{z_2}{1!}+\frac{z_1^2}{2!}\right)+\cdots. \qquad (19)$$

Il termine generale della serie è

$$\frac{z_2^n}{n!}+\frac{z_2^{n-1}}{(n-1)!}\frac{z_1}{1!}+\frac{z_2^{n-2}}{(n-2)!}\frac{z_1^2}{2!}+\cdots+\frac{z_1^n}{n!}$$
$$=\frac{1}{n!}\left(z_2^n+\frac{n!}{1!(n-1)!}z_2^{n-1}z_1+\frac{n!}{2!(n-2)!}z_2^{n-2}z_1^2+\cdots+z_1^n\right).$$

Applicando la formula del binomio di Newton, esso assume la forma

$$\frac{(z_1+z_2)^n}{n!}.$$

Il termine generale della (19) coincide quindi con quello della serie di $e^{z_1+z_2}$, e la regola di moltiplicazione (18) è dimostrata. Il riordinamento è lecito per la convergenza assoluta delle serie.

Il teorema di moltiplicazione e la formula di Eulero permettono di esprimere e^z in forma finita, senza serie, mediante funzioni reali. Ponendo

$$z=x+iy,$$

otteniamo

$$e^z=e^{x+iy}=e^x\cdot e^{iy},$$

e, poiché

$$e^{iy}=\cos y+i\sin y,$$

troviamo

$$e^z = e^x(\cos y + i \sin y). \qquad (20)$$

La formula è particolarmente comoda per studiare le proprietà di e^z. Ne ricordiamo due: (1) la funzione non si annulla mai, poiché $e^x \neq 0$ e cos y e sin y non si annullano contemporaneamente; (2) la funzione ha periodo $2\pi i$, ossia

$$e^{z+2\pi i} = e^z.$$

Quest'ultima affermazione segue dal teorema di moltiplicazione e dall'uguaglianza

$$e^{2\pi i} = \cos 2\pi + i \sin 2\pi = 1.$$

Le formule (17) permettono di studiare cos z e sin z nel campo complesso. Lasciamo al lettore la verifica che entrambe hanno periodo 2π e che continuano a valere le formule di addizione del seno e del coseno.

Il concetto generale di funzione complessa e la derivabilità.

Le serie di potenze permettono di definire funzioni analitiche complesse. È però interessante studiare le operazioni fondamentali dell'analisi, e in particolare la derivazione, per una funzione complessa arbitraria. Emergono allora fatti profondi. Da una parte, come vedremo, una funzione che possiede la derivata prima in tutti i punti di un intorno di z_0 possiede necessariamente in z_0 derivate di ogni ordine ed è sviluppabile in una serie di potenze centrata in quel punto: è dunque analitica. Considerando le funzioni complesse derivabili si ritorna quindi immediatamente alla classe delle funzioni analitiche. D'altra parte, lo studio della derivata rivela il comportamento geometrico delle funzioni complesse e i loro legami con i problemi della fisica matematica.

Nel seguito chiameremo perciò analitica in z_0 una funzione che possiede la derivata in tutti i punti di un intorno di z_0.

Seguendo la definizione generale di funzione, diremo che una variabile complessa w è funzione della variabile complessa z se esiste una legge che consente di trovare il valore di w conoscendo quello di z.

Ogni numero $z = x + iy$ è rappresentato da un punto (x, y) del piano Oxy; rappresentiamo analogamente $w = u + iv$ in un piano Ouv, il piano dei valori della funzione. Geometricamente, $w = f(z)$ stabilisce una corrispondenza fra punti del piano dell'argomento e punti del piano dei valori. In altre parole, una funzione complessa determina una trasformazione da un piano all'altro. Assegnarla significa stabilire la corrispondenza fra le coppie (x, y) e (u, v), cioè assegnare due funzioni

$$u = \phi(x, y), \qquad v = \psi(x, y),$$

per le quali, evidentemente,

$$w = u + iv = \phi(x, y) + i\psi(x, y).$$

Per esempio, se

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

allora

$$u = \phi(x, y) = x^2 - y^2, \qquad v = \psi(x, y) = 2xy.$$

La derivata di una funzione complessa si definisce formalmente come nel caso reale: è il limite del rapporto incrementale,

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}, \tag{21}$$

se tale limite esiste.

L'esistenza delle derivate parziali rispetto a x e y delle due funzioni reali u e v che compongono $w = f(z)$ non basta a garantire l'esistenza di $f'(z)$. In generale il limite del rapporto incrementale dipende dalla direzione lungo la quale $z' = z + \Delta z$ si avvicina a z (figura 3). Perché la derivata esista, il limite deve essere indipendente da ogni modo di avvicinamento. Consideriamo, per esempio, un avvicinamento parallelo all'asse Ox o all'asse Oy.

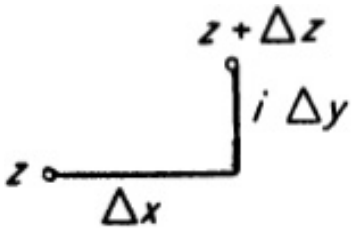


Figura 3.

Nel primo caso

$$\Delta z = \Delta x,$$
$$f(z + \Delta z) - f(z) = u(x + \Delta x, y) - u(x, y) + i[v(x + \Delta x, y) - v(x, y)],$$

e il rapporto incrementale

$$\frac{f(z + \Delta z) - f(z)}{\Delta z} = \frac{u(x + \Delta x, y) - u(x, y)}{\Delta x} + i\,\frac{v(x + \Delta x, y) - v(x, y)}{\Delta x}$$

tende, per $\Delta x \to 0$, a

$$\frac{\partial u}{\partial x} + i\,\frac{\partial v}{\partial x}. \tag{22}$$

Nel secondo caso

$$\Delta z = i\,\Delta y,$$

e il rapporto incrementale

$$\frac{f(z+\Delta z)-f(z)}{\Delta z} = -i\,\frac{u(x,y+\Delta y)-u(x,y)}{\Delta y} + \frac{v(x,y+\Delta y)-v(x,y)}{\Delta y}$$

conduce, al limite, a

$$\partial v/\partial y - i\,\partial u/\partial y. \quad (23)$$

Se $w = f(z)$ possiede la derivata, queste due espressioni devono essere uguali; dunque

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y}, \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}. \end{aligned} \qquad (24)$$

Queste equazioni sono condizioni necessarie per l'esistenza della derivata di $w = u + iv$. Si può dimostrare che sono anche sufficienti, purché u e v siano differenziabili come funzioni di due variabili reali. Non dimostreremo qui la sufficienza delle condizioni (24), note come equazioni di Cauchy-Riemann.

È facile verificare che le consuete regole di derivazione delle funzioni reali si trasferiscono senza modifiche alle funzioni complesse. Ciò vale per la derivata di z^n e per quelle di una somma, di un prodotto o di un quoziente. Il metodo di dimostrazione è esattamente lo stesso, sostituendo alle grandezze reali quelle complesse. Ne segue che ogni polinomio in z,

$$w = a_0 + a_1 z + \cdots + a_n z^n$$

è derivabile ovunque. Ogni funzione razionale, quoziente di due polinomi,

$$w = \frac{a_0 + a_1 z + \cdots + a_n z^n}{b_0 + b_1 z + \cdots + b_m z^m}$$

è derivabile in tutti i punti in cui il denominatore non è nullo.

Per stabilire la derivabilità di $w = e^z$ possiamo usare le condizioni di Cauchy-Riemann. Dalla (20), in questo caso,

$$u = e^x \cos y, \qquad v = e^x \sin y;$$

Sostituendo queste funzioni nella (24), si verifica che le equazioni di Cauchy-Riemann sono soddisfatte. La derivata si calcola, per esempio, con la (22), ottenendo

$$\frac{dw}{dz} = e^z.$$

Mediante la (17) si dimostrano facilmente la derivabilità delle funzioni trigonometriche e la validità delle consuete formule dell'analisi per le loro derivate.

La funzione Ln z. Non esamineremo tutte le funzioni elementari complesse. Per i nostri scopi è però importante conoscere alcune proprietà di Ln z. Come nel campo reale, poniamo

$$w = \operatorname{Ln} z,$$

se

$$z = e^w.$$

Per studiare Ln z, scriviamo z in forma trigonometrica:

$$z = r(\cos\phi + i\sin\phi).$$

Applicando a e^w il teorema di moltiplicazione, otteniamo

$$z = e^w = e^{u+iv} = e^u e^{iv} = e^u(\cos v + i\sin v).$$

Uguagliando le due espressioni di z, si ha

$$e^u = r, \qquad (\alpha)$$

$$\cos v + i\sin v = \cos\phi + i\sin\phi. \qquad (\beta)$$

Poiché u e r sono reali, dalla (α) ricaviamo

$$u = \ln r,$$

dove ln r è il consueto logaritmo naturale di un numero reale positivo. L'equazione (β) può essere soddisfatta soltanto se

$$\cos v = \cos\phi, \quad \sin v = \sin\phi,$$

e in tal caso v e ϕ differiscono per un multiplo intero di 2π:

$$v = \phi + 2\pi n,$$

Ogni intero n fornisce una soluzione della (β). Dalle espressioni di u e v ricaviamo

$$\operatorname{Ln} z = \ln r + i(\phi + 2\pi n). \qquad (25)$$

La formula (25) definisce Ln z per ogni numero complesso diverso da zero. Essa estende il logaritmo dai numeri positivi anche ai numeri negativi e complessi.

L'espressione contiene un intero arbitrario n: Ln z è dunque una funzione a più valori. Per ogni n otteniamo uno dei suoi possibili valori; fissando n e una determinazione dell'argomento, ne scegliamo uno.

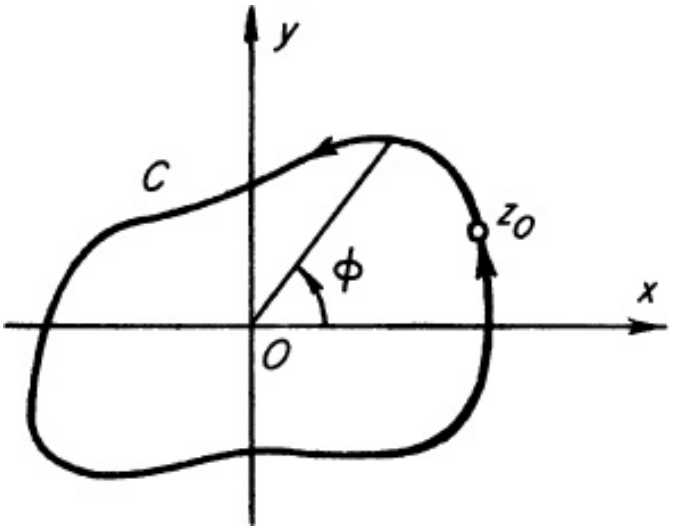


Figura 4.

I diversi valori di Ln z sono però strettamente collegati. Fissiamo, per esempio, $n = 0$ in un punto z_0 e facciamo percorrere a z, in senso antiorario, una curva semplice chiusa C che circonda l'origine e ritorna a z_0 (figura 4). Seguendo continuamente l'argomento ϕ, al termine del giro esso aumenta complessivamente di 2π. Fissato quindi il valore iniziale

$$(\mathrm{Ln}\ z)_0 = \ln r_0 + i\phi_0$$

e variandolo con continuità mentre z percorre la curva, ritorniamo a z_0 con un altro valore della funzione:

$$(\mathrm{Ln}\ z)_0 = \ln r_0 + i(\phi_0 + 2\pi).$$

Possiamo dunque passare con continuità da un valore di Ln z a un altro: basta far compiere al punto un numero sufficiente di giri intorno all'origine. Il punto $z = 0$ si chiama punto di diramazione di Ln z.

Se vogliamo considerare una sola determinazione di Ln z, dobbiamo impedire a z di percorrere una curva chiusa che circondi l'origine. Possiamo farlo tracciando una curva semplice dall'origine all'infinito e vietandone l'attraversamento: questa curva si chiama taglio. Nel piano tagliato, il prolungamento continuo non porta a valori diversi dello stesso logaritmo nello stesso punto. Partendo da un valore assegnato in z_0, otteniamo così un solo valore in ogni punto. I valori scelti costituiscono un ramo univoco della funzione.

Se, per esempio, il taglio coincide con il semiasse reale negativo, si ottiene un ramo univoco restringendo l'argomento all'intervallo

$$(2k - 1)\pi < \phi < (2k + 1)\pi,$$

dove k è un intero arbitrario.

Scelto un ramo univoco del logaritmo, possiamo studiarne la derivabilità. Ponendo

$$r = \sqrt{x^2 + y^2}, \qquad \phi = \arctan \frac{y}{x},$$

L'argomento va scelto nel quadrante corretto e nel ramo fissato; la formula con arctan vale localmente con tale scelta.

è facile verificare le condizioni di Cauchy-Riemann. La derivata, calcolata per esempio mediante la (22), è

$$\frac{d \operatorname{Ln} z}{dz} = \frac{1}{z}.$$

Sottolineiamo che la derivata di Ln *z* è univoca: è la stessa per tutti i rami.

§2. Le funzioni complesse e i problemi della fisica matematica

Il legame con i problemi dell'idrodinamica.

Le condizioni di Cauchy-Riemann collegano la teoria delle funzioni complesse ai problemi della fisica matematica. Illustreremo questo legame con alcuni problemi d'idrodinamica.

Fra tutti i moti possibili di un fluido, hanno grande importanza quelli stazionari, nei quali la distribuzione spaziale delle velocità non cambia nel tempo. Un osservatore su un ponte che guarda il fiume scorrere intorno a un pilone può vedere, per esempio, un moto approssimativamente stazionario. Talvolta il moto è stazionario per un osservatore che si muove su un veicolo. Il flusso intorno a una nave in avanzamento può apparire non stazionario a un osservatore a terra e stazionario a uno a bordo, nelle opportune condizioni. Analogamente, per un passeggero di un aeroplano che vola a velocità costante, il flusso d'aria perturbato dal velivolo può essere stazionario.

In un moto stazionario, il vettore velocità *V* della particella che passa per un dato punto dello spazio non cambia nel tempo. Se la stazionarietà vale per un osservatore in moto, il vettore non cambia nei punti con coordinate costanti nel sistema di riferimento solidale con l'osservatore.

Un'altra classe importante è quella dei moti piani paralleli: le velocità delle particelle sono ovunque parallele a un piano e la loro distribuzione è identica su tutti i piani paralleli a esso.

Immaginiamo una massa fluida infinitamente estesa che scorra intorno a un corpo cilindrico, perpendicolarmente alle generatrici. La distribuzione delle velocità sarà uguale in tutti i piani perpendicolari alle generatrici, e il moto sarà piano parallelo. In molti casi questa è una buona approssimazione. Per esempio, in una sezione perpendicolare all'apertura alare, il moto dell'aria può considerarsi approssimativamente piano parallelo, purché la sezione non sia troppo vicina alla fusoliera o all'estremità dell'ala.

Mostreremo come applicare la teoria delle funzioni complesse allo studio di un moto stazionario piano parallelo.

Supporremo che il fluido sia incomprimibile, cioè che la sua densità non vari con la pressione. Questa ipotesi è una buona approssimazione, per esempio, per l'acqua. Anche l'aria può essere trattata come incomprimibile quando la velocità è sufficientemente piccola. Il testo originale indicava come limite orientativo da 0,6 a 0,8 volte la velocità del suono, assunta pari a 330 m/s: questa indicazione storica non costituisce una soglia generale di validità dell'approssimazione.

Il flusso di un fluido è caratterizzato dalla distribuzione delle velocità delle sue particelle. Se il moto è piano parallelo, basta determinare le velocità in uno dei piani paralleli ai quali esso si svolge.

Indichiamo con $V(x, y, t)$ la velocità della particella che passa per (x, y) all'istante t. Nel moto stazionario V non dipende dal tempo. Il vettore è individuato dalle componenti u, v lungo gli assi. Consideriamo le traiettorie delle particelle: nel moto stazionario le velocità delle particelle che passano successivamente per un dato punto non cambiano. Conoscendo il campo delle velocità, cioè le componenti in funzione di x e y, possiamo determinare le traiettorie osservando che la velocità è sempre tangente a esse. Si ottiene

$$\frac{dy}{dx} = \frac{v(x, y)}{u(x, y)} .$$

Questa è l'equazione differenziale delle traiettorie. Nel moto stazionario esse coincidono con le linee di corrente. Per ogni punto in cui la velocità non è nulla passa una sola linea di corrente, sotto le usuali ipotesi di regolarità.

Un ruolo importante è svolto dalla funzione di corrente. Fissata una linea C_0, consideriamo un canale immaginario delimitato da quattro pareti: la superficie cilindrica, con generatrici perpendicolari al piano del moto, passante per C_0; l'analoga superficie passante per una linea vicina C_1; il piano del moto; e un piano parallelo a distanza unitaria (figura 5). La quantità di fluido che attraversa in un'unità di tempo due sezioni arbitrarie γ_1 e γ_2 del canale è la stessa. Infatti, per la densità costante, la quantità contenuta fra le due sezioni non può cambiare; inoltre le pareti laterali sono formate da linee di corrente e non vengono attraversate dal fluido. La quantità che esce attraverso una sezione deve quindi uguagliare quella che entra attraverso l'altra.

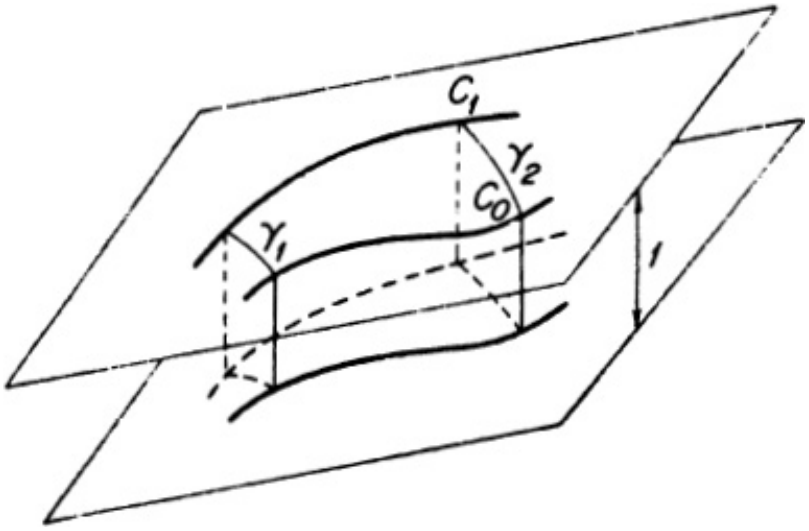


Figura 5.

Chiamiamo funzione di corrente la funzione $\psi(x, y)$ che assume sulla linea C_1 un valore costante uguale alla portata, con il segno stabilito dall'orientazione, attraverso la sezione del canale costruito sulle curve C_0 e C_1.

La funzione di corrente è determinata a meno di una costante additiva, dipendente dalla scelta della linea iniziale C_0. Conoscendola, le linee di corrente sono evidentemente date da

$$\psi(x, y) = \text{const.}$$

Esprimiamo ora le componenti della velocità in un punto $M(x, y)$ mediante le derivate della funzione di corrente. Consideriamo il canale formato dalla linea C passante per M e dalla linea vicina C' passante per $M'(x, y + \Delta y)$, insieme con due piani paralleli al moto e distanti un'unità. Calcoliamo la quantità di fluido q che attraversa la sezione MM' nel tempo dt.

Da una parte, per definizione di funzione di corrente,

$$q = (\psi' - \psi)\, dt.$$

D'altra parte, q è uguale al volume del solido ottenuto riportando il vettore $V\,dt$ da ogni punto della sezione MM' (figura 6). Se MM' è piccolo, possiamo considerare V costante su tutta la sezione e uguale al valore in M. La base del parallelepipedo ha area $\Delta y \times 1$; lo spessore unitario non è indicato nella figura. L'altezza è la proiezione di $V\,dt$ sull'asse Ox, cioè $u\,dt$. Pertanto, al primo ordine,

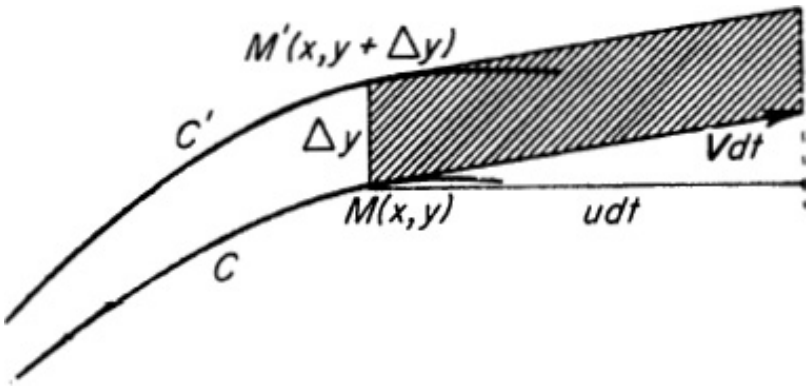


Figura 6.

$$q \approx u\,\Delta y\,dt$$

e dunque

$$u\,\Delta y \approx \Delta\psi.$$

Dividendo per Δy e passando al limite, otteniamo

$$u = \frac{\partial\psi}{\partial y}. \qquad (26)$$

Un ragionamento analogo fornisce la seconda componente:

$$v = -\frac{\partial\psi}{\partial x}. \qquad (26')$$

Per descrivere il campo delle velocità introduciamo, accanto alla funzione di corrente, un'altra funzione, collegata alla rotazione di piccoli elementi di fluido. Se immaginassimo di solidificare un elemento, esso avrebbe in generale anche un moto di rotazione. Se però il fluido parte dalla quiete, non ha attrito interno e sono soddisfatte le consuete ipotesi sulle forze agenti, si dimostra che tale rotazione non si genera. I moti privi di questa rotazione si chiamano irrotazionali e hanno un ruolo fondamentale nello studio dei corpi in moto in un fluido. L'idromeccanica mostra che per un flusso irrotazionale esiste, almeno localmente, una funzione $\phi(x, y)$ tale che

$$u = \frac{\partial\phi}{\partial x}, \qquad v = \frac{\partial\phi}{\partial y}; \qquad (27)$$

La funzione ϕ si chiama potenziale delle velocità. Nel seguito considereremo moti dotati di questo potenziale.

Confrontando le espressioni delle componenti della velocità ottenute dalla funzione di corrente e dal potenziale, si ricava il risultato notevole seguente.

Il potenziale delle velocità $\phi(x, y)$ e la funzione di corrente $\psi(x, y)$ di un flusso incomprimibile e irrotazionale soddisfano le equazioni di Cauchy-Riemann:

$$\frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y}, \qquad \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x}. \tag{28}$$

In altre parole, la funzione complessa

$$w = \phi(x, y) + i\psi(x, y)$$

è derivabile nel senso complesso. Viceversa, le parti reale e immaginaria di una funzione analitica soddisfano le condizioni di Cauchy-Riemann e possono essere interpretate come potenziale delle velocità e funzione di corrente di un fluido incomprimibile e irrotazionale. La funzione w si chiama funzione caratteristica del flusso, o potenziale complesso.

Consideriamo il significato della derivata di w. Usando, per esempio, la (22), abbiamo

$$\frac{dw}{dz} = \frac{\partial \phi}{\partial x} + i\frac{\partial \psi}{\partial x}.$$

Dalle (27) e (26′) si trova

$$\frac{dw}{dz} = u - iv$$

oppure, passando ai complessi coniugati,

$$u + iv = \left(\overline{\frac{dw}{dz}}\right), \tag{29}$$

dove la barra sopra dw/dz indica il complesso coniugato.

Il numero complesso che rappresenta il vettore velocità è quindi il coniugato della derivata della funzione caratteristica.

Esempi di moti piani paralleli.

Consideriamo alcuni esempi. Sia

$$w = Az, \tag{30}$$

dove A è una costante complessa. Dalla (29) segue

$$u + iv = \operatorname{Re} A - i \operatorname{Im} A.$$

La funzione lineare (30) descrive dunque un flusso con vettore velocità costante. Ponendo

$$A = u_0 - iv_0 ,$$

e separando parte reale e immaginaria di w, si ha

$$\phi(x, y) = u_0 x + v_0 y,$$
$$\psi(x, y) = u_0 y - v_0 x,$$

Le linee di corrente sono quindi rette parallele al vettore velocità (figura 7).

Come secondo esempio consideriamo

$$w = Az^2,$$

con A reale. Per rappresentare il flusso determiniamo prima le linee di corrente. In questo caso

$$\psi(x, y) = 2Axy,$$

e le loro equazioni sono

$$xy = \text{const.}$$

Sono iperboli aventi come asintoti gli assi coordinati (figura 8). Le frecce indicano il verso del moto per $A > 0$. Anche gli assi Ox e Oy appartengono alle linee di corrente, con un punto di ristagno nell'origine.

Nel modello privo di viscosità possiamo sostituire una linea di corrente con una parete rigida senza modificare il resto del flusso, poiché il fluido scorre tangenzialmente alla parete. Costruendo pareti lungo i semiassi positivi, indicati con linee più spesse nella figura 8, otteniamo lo schema di un moto irrotazionale in un angolo.

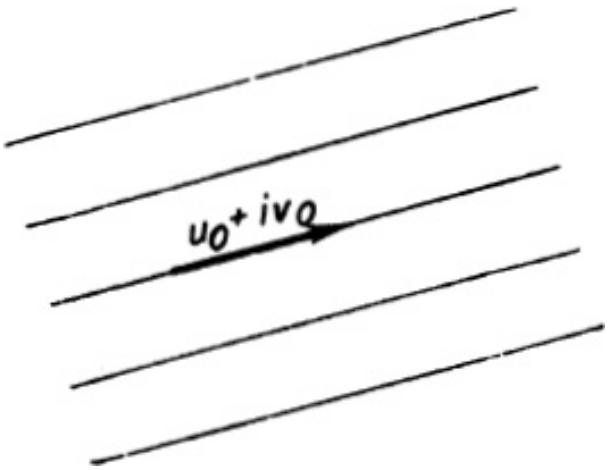


Figura 7.

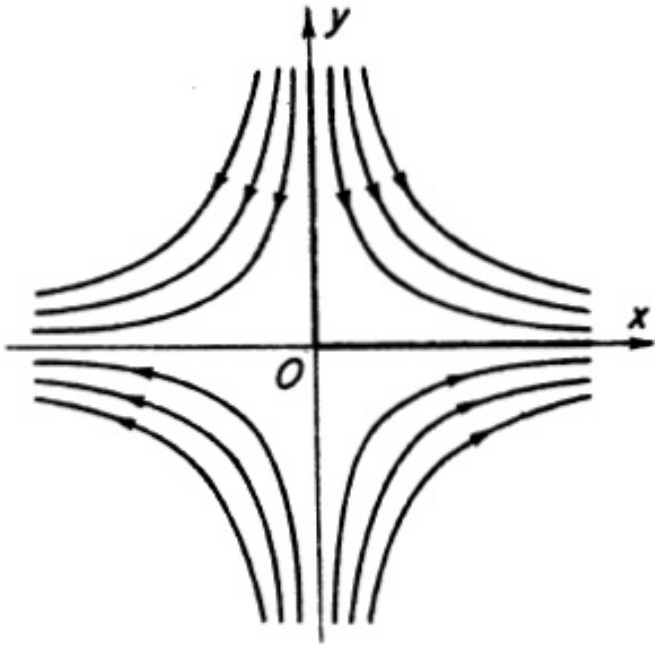


Figura 8.

Un esempio importante è dato dalla funzione

$$w = a\left(z + \frac{R^2}{z}\right), \tag{31}$$

dove a e R sono reali positivi.

La funzione di corrente è

$$\psi = a\left(y - \frac{R^2 y}{x^2 + y^2}\right),$$

e l'equazione delle linee di corrente è quindi

$$y - \frac{R^2 y}{x^2 + y^2} = \text{const.}$$

Ponendo la costante uguale a zero, otteniamo $y = 0$ oppure $x^2 + y^2 = R^2$. La circonferenza di raggio R è dunque una linea di corrente. Sostituendone l'interno con un corpo solido, otteniamo il flusso intorno a un cilindro circolare. La figura 9 ne mostra le linee di corrente. Dalla (29) la velocità è

$$u + iv = a\left(1 - \frac{R^2}{\bar{z}^2}\right).$$

A grande distanza dal cilindro si ha

$$\lim_{z \to \infty} (u + iv) = a;$$

Lontano dal cilindro la velocità tende quindi a un valore costante e il flusso diventa uniforme. La funzione (31) descrive dunque il moto intorno a un cilindro circolare di un fluido che, a grande distanza, scorre uniformemente.

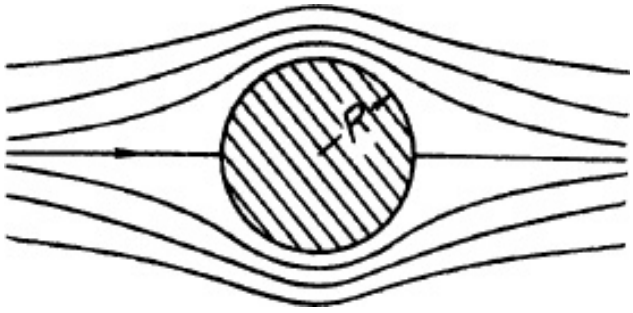


Figura 9.

Le idee fondamentali della teoria dell'ala; il teorema di Žukovskij.

L'applicazione delle funzioni complesse ai moti piani paralleli condusse Žukovskij e Čaplygin a importanti scoperte aerodinamiche. Lo studio del flusso intorno ai corpi li portò alla legge della portanza dell'ala. Per presentare le idee alla base di questa scoperta, consideriamo un altro esempio concreto, con funzione caratteristica

$$w = \frac{\Gamma}{2\pi i} \operatorname{Ln} z,$$

dove Γ è una costante reale. Benché w abbia più valori, la sua derivata

$$\frac{dw}{dz} = \frac{\Gamma}{2\pi i} \frac{1}{z} \qquad (32)$$

è univoca, e quindi il campo delle velocità è determinato senza ambiguità. Ponendo $z = re^{i\theta}$, dalla (25) otteniamo il potenziale e la funzione di corrente:

$$\phi = \frac{\Gamma}{2\pi} \theta, \qquad \psi = -\frac{\Gamma}{2\pi} \ln r.$$

La seconda formula mostra che le linee di corrente sono circonferenze r = costante (figura 10).

La velocità si ricava dalla (29):

$$u + iv = -\frac{\Gamma}{2\pi i} \frac{1}{\bar{z}}.$$

In particolare, il suo modulo è

$$V = |u + iv| = \frac{|\Gamma|}{2\pi} \frac{1}{r},$$

La velocità ha quindi modulo costante su ciascuna linea di corrente. Un esame più dettagliato mostra che il moto avviene in senso antiorario per $\Gamma > 0$ e orario per $\Gamma < 0$.

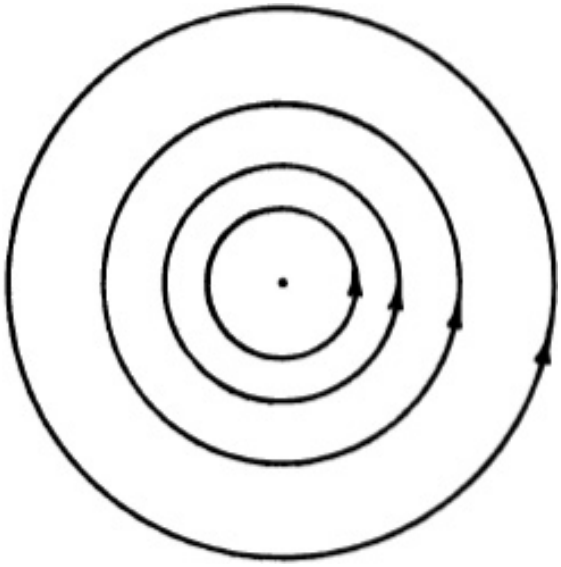


Figura 10.

Sostituendo una linea di corrente con una parete rigida, otteniamo un moto circolare del fluido intorno a un cilindro. Un tale moto si chiama circolatorio.

Il potenziale di questo moto non è univoco. Dopo un giro antiorario lungo un contorno chiuso intorno al cilindro cambia di Γ. Tale variazione del potenziale è la circolazione del flusso.

Sommando alla funzione caratteristica del flusso intorno al cilindro (31) quella di un moto circolatorio, qui scelto orario per $\Gamma > 0$, otteniamo

$$w = a\left(z + \frac{R^2}{z}\right) - \frac{\Gamma}{2\pi i}\,\mathrm{Ln}\, z. \qquad (33)$$

Anche questa funzione descrive un flusso intorno a un cilindro di raggio R. Infatti, sulla circonferenza di raggio R la funzione di corrente è costante, poiché sono costanti le parti immaginarie di entrambi i termini. La velocità definita dalla (33) tende ancora ad a per $z \to \infty$. Per ogni Γ, la (33) rappresenta dunque un flusso uniforme all'infinito che aggira il cilindro. La figura 11 illustra il caso $\Gamma > 0$. Il flusso non è simmetrico rispetto all'asse orizzontale: i punti di ristagno a e b, nei quali le correnti si separano o si ricongiungono, sono spostati verso il basso, nel caso rappresentato. Il potenziale è a più valori e dopo un giro cambia di una quantità di modulo $|\Gamma|$, con segno dipendente dall'orientazione.

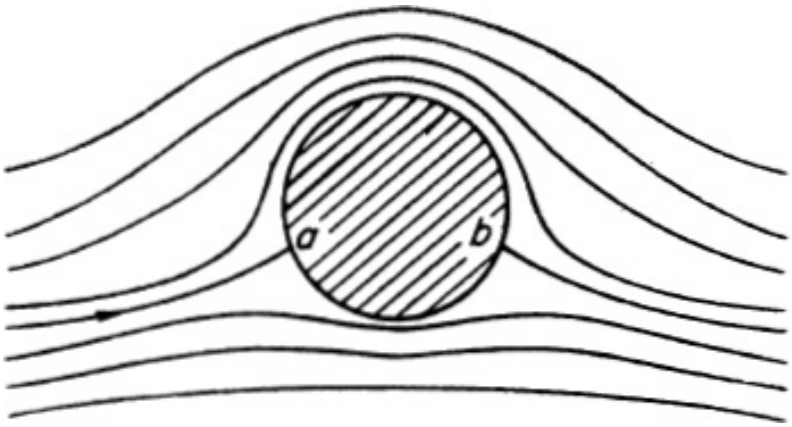


Figura 11.

Per ragioni di simmetria, il flusso intorno a un cilindro si considera spesso senza circolazione, nella forma (31). Intorno a corpi non simmetrici si possono invece stabilire moti con potenziale a più valori. Discuteremo fra poco il significato fisico di questo fatto. I metodi delle funzioni complesse permettono di determinare flussi intorno a corpi di forma arbitraria: nella prossima sezione vedremo come, partendo dal cilindro, costruire flussi con potenziale univoco o multivoco intorno ad altri profili.

Nel caso di un'ala si considera un profilo con un bordo posteriore affilato: il profilo si assottiglia verso il bordo d'uscita. Costruendo un flusso con potenziale univoco, il punto di ristagno presso il quale la corrente lascia il profilo non coincide, in generale, con il bordo (figura 12a). Questo flusso presenta una difficoltà fisica: al bordo affilato il modello prevede velocità infinita e una corrispondente caduta illimitata della pressione. La soluzione fisicamente selezionata è quella in cui la corrente lascia regolarmente il bordo d'uscita (figura 12b). Essa ha in generale un potenziale a più valori, cioè una circolazione non nulla. Anche qui la circolazione si definisce come variazione del potenziale lungo un giro completo intorno al profilo, con orientazione fissata.



Figura 12.

La condizione di realizzabilità fisica per cui il flusso lascia il bordo d'uscita è chiamata nell'opera postulato di Čaplygin.

La scoperta notevole di Žukovskij consiste nel fatto che la circolazione genera sull'ala una forza portante per unità di apertura, perpendicolare alla velocità a della corrente incidente, data nella convenzione di segno adottata da

$$\rho a \Gamma,$$

dove ρ è la densità del mezzo e Γ la circolazione; il modulo della forza è $\rho a|\Gamma|$ (figura 13).



Figura 13.

Il teorema di Žukovskij sulla portanza è uno dei fondamenti dell'aerodinamica. Non ne daremo la dimostrazione; ricordiamo soltanto che le dimostrazioni usuali si basano sugli integrali di funzioni complesse.

I risultati fondamentali di Žukovskij e Čaplygin furono ampiamente sviluppati dagli scienziati sovietici.

Applicazioni ad altri problemi della fisica matematica.

La teoria delle funzioni complesse trova applicazioni non soltanto nella teoria dell'ala, ma anche in molti altri problemi d'idrodinamica.

Il suo campo d'applicazione è però assai più vasto e comprende numerosi altri problemi della fisica matematica. Per illustrarlo, ritorniamo alle condizioni di Cauchy-Riemann:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y},$$
$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

Ne ricaviamo un'equazione soddisfatta dalla parte reale di una funzione analitica. Derivando la prima rispetto a x, la seconda rispetto a y e sommando, otteniamo

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Questa è l'equazione di Laplace, già incontrata nel capitolo VI. Moltissimi problemi della fisica e della meccanica conducono a essa. Per esempio, in condizioni di equilibrio termico, in un mezzo omogeneo senza sorgenti, la temperatura soddisfa l'equazione di Laplace. La stessa equazione compare nello studio dei campi magnetici ed elettrostatici, nelle regioni appropriate, e nella filtrazione di liquidi attraverso mezzi porosi. Nei problemi piani di questo tipo i metodi delle funzioni complesse trovano ampie applicazioni.

Non soltanto l'equazione di Laplace, ma anche equazioni più generali della fisica matematica possono essere collegate alle funzioni complesse. Un esempio importante è costituito dai problemi piani dell'elasticità. Le basi di questa applicazione furono poste dagli scienziati sovietici G. B. Kolosov e N. I. Muschelišvili.

§3. Le funzioni complesse e la geometria

Proprietà geometriche delle funzioni derivabili.

Come per le funzioni reali, l'interpretazione geometrica ha grande importanza nella teoria delle funzioni analitiche complesse. Le proprietà geometriche offrono una rappresentazione intuitiva di quelle analitiche e hanno inoltre dato origine a un insieme autonomo di problemi, la teoria geometrica delle funzioni. Come abbiamo detto, una funzione $w = f(z)$ è geometricamente una trasformazione dal piano z al piano w. Essa si può descrivere mediante due funzioni di due variabili reali:

$$u = u(x, y),$$
$$v = v(x, y).$$

Per studiare la trasformazione in un intorno molto piccolo di un punto, possiamo sviluppare queste funzioni intorno al punto e limitarci ai termini di primo ordine:

$$u - u_0 = \left(\frac{\partial u}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial u}{\partial y}\right)_0 (y - y_0) + \cdots,$$
$$v - v_0 = \left(\frac{\partial v}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial v}{\partial y}\right)_0 (y - y_0) + \cdots,$$

Le derivate sono calcolate in (x_0, y_0). Nell'intorno di un punto, una trasformazione sufficientemente regolare può dunque essere approssimata da una trasformazione affine,[3]

$$u - u_0 = a(x - x_0) + b(y - y_0),$$
$$v - v_0 = c(x - x_0) + d(y - y_0),$$

dove

$$a = \left(\frac{\partial u}{\partial x}\right)_0, \qquad b = \left(\frac{\partial u}{\partial y}\right),$$
$$c = \left(\frac{\partial v}{\partial x}\right)_0, \qquad d = \left(\frac{\partial v}{\partial y}\right)_0.$$

Esaminiamo la trasformazione prodotta da una funzione analitica vicino a $z = x + iy$. Sia C una curva uscente da z; i punti corrispondenti nel piano w descrivono una curva Γ uscente da w. Se z' è un punto vicino e w' il suo corrispondente, per $z' \to z$ si ha $w' \to w$ e

$$\frac{w' - w}{z' - z} \to f'(z). \qquad (34)$$

In particolare,

$$\frac{|w' - w|}{|z' - z|} \to |f'(z)| . \qquad (35)$$

Possiamo esprimere il risultato nel modo seguente.

Nel punto z, il limite del rapporto fra le lunghezze delle corde corrispondenti nel piano w e nel piano z è lo stesso per tutte le curve uscenti da z. Equivalentemente, il rapporto fra elementi lineari corrispondenti nei due piani non dipende dalla curva considerata.

La quantità $|f'(z)|$, che misura l'ingrandimento degli elementi lineari, si chiama fattore di dilatazione nel punto z.

Supponiamo ora che $f'(z) \neq 0$, cosicché il suo argomento sia determinato a meno di multipli di 2π.[4] Calcoliamolo usando la (34):

$$\arg \frac{w' - w}{z' - z} = \arg (w' - w) - \arg (z' - z),$$

L'argomento di $w' - w$ è l'angolo β' fra la corda ww' e l'asse reale, mentre l'argomento di $z' - z$ è l'angolo α' fra zz' e l'asse reale. Indichiamo con α e β gli angoli corrispondenti delle tangenti a C e Γ nei punti z e w (figura 14). Per $z' \to z$,

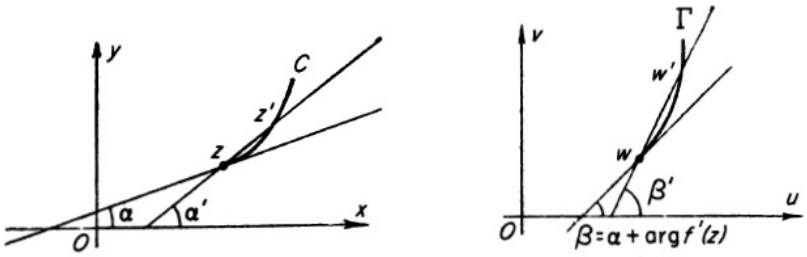


Figura 14.

$$\alpha' \to \alpha, \quad \beta' \to \beta,$$

e al limite otteniamo

$$\arg f'(z) = \beta - \alpha. \qquad (36)$$

L'argomento di $f'(z)$ è dunque l'angolo ϕ di cui bisogna ruotare la tangente a C in z perché assuma la direzione della tangente a Γ in w. Per questo si chiama angolo di rotazione della trasformazione nel punto z.

Dalla (36) il lettore può ricavare facilmente le proposizioni seguenti.

Nel passaggio dal piano z al piano w, le tangenti a tutte le curve uscenti da un punto ruotano dello stesso angolo.

Se C_1 e C_2 escono da z, e Γ_1 e Γ_2 sono le loro immagini uscenti da w, l'angolo fra le curve corrispondenti è uguale nei due piani.

In una trasformazione analitica, in ogni punto in cui $f'(z) \neq 0$ tutti gli elementi lineari sono quindi moltiplicati per lo stesso fattore e gli angoli orientati fra direzioni corrispondenti restano invariati.

Le trasformazioni con queste proprietà si chiamano conformi.

Le proprietà appena dimostrate fanno supporre che, vicino a un punto z_0 con $f'(z_0) \neq 0$, la trasformazione sia biunivoca: a ciascun z corrisponde un solo w e, viceversa, a ciascun w dell'immagine corrisponde un solo z. Questa proprietà locale può essere dimostrata rigorosamente.

Per chiarire la differenza fra trasformazioni conformi e altri tipi di trasformazioni, consideriamone una arbitraria vicino a un punto. Prendendo i primi termini degli sviluppi di u e v, otteniamo

$$u - u_0 = \left(\frac{\partial u}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial u}{\partial y}\right)_0 (y - y_0) + \cdots,$$

$$v - v_0 = \left(\frac{\partial v}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial v}{\partial y}\right)_0 (y - y_0) + \cdots.$$

Trascurando i termini d'ordine superiore in un piccolo intorno di (x_0, y_0), la trasformazione si comporta come una trasformazione affine. La sua parte lineare è invertibile se il determinante non si annulla:

$$\Delta = \left(\frac{\partial u}{\partial x}\right)_0 \left(\frac{\partial v}{\partial y}\right)_0 - \left(\frac{\partial u}{\partial y}\right)_0 \left(\frac{\partial v}{\partial x}\right)_0 \neq 0.$$

Se $\Delta = 0$, per descrivere il comportamento vicino al punto occorre considerare anche termini d'ordine superiore.[5]

Se $u + iv$ è analitica, le equazioni di Cauchy-Riemann permettono di esprimere le derivate rispetto a y mediante quelle rispetto a x. Si ottiene

$$\Delta = \left(\frac{\partial u}{\partial x}\right)_0^2 + \left(\frac{\partial v}{\partial x}\right)_0^2 = \left|\left(\frac{\partial u}{\partial x}\right)_0 + i\left(\frac{\partial v}{\partial x}\right)_0\right|^2 = |f'(z_0)|^2,$$

La trasformazione è dunque localmente invertibile quando $f'(z_0) \neq 0$. Ponendo $f'(z_0) = r(\cos\phi + i \sin\phi)$, si ha

$$\left(\frac{\partial u}{\partial x}\right)_0 = \left(\frac{\partial v}{\partial y}\right)_0 = r\cos\phi,$$

$$\left(\frac{\partial u}{\partial y}\right)_0 = -\left(\frac{\partial v}{\partial x}\right)_0 = -r\sin\phi,$$

e la trasformazione vicino a (x_0, y_0) assume la forma

$$u - u_0 = r[(x - x_0)\cos\phi - (y - y_0)\sin\phi] + \cdots,$$
$$v - v_0 = r[(x - x_0)\sin\phi + (y - y_0)\cos\phi] + \cdots.$$

Le formule mostrano che, al primo ordine, una trasformazione analitica consiste in una rotazione di angolo ϕ e in una dilatazione di fattore r, oltre alla traslazione che fa corrispondere i punti di riferimento. Le espressioni fra parentesi sono infatti le note formule della rotazione piana; la moltiplicazione per r produce la dilatazione.

Per capire che cosa possa accadere quando $f'(z) = 0$, consideriamo

$$w = z^n. \qquad (37)$$

con n intero maggiore di uno. La derivata $w' = nz^{n-1}$ si annulla nell'origine. Conviene studiare la (37) mediante le coordinate polari, cioè la forma trigonometrica dei numeri complessi. Sia

$$z = r(\cos\phi + i\sin\phi),$$
$$w = \rho(\cos\theta + i\sin\theta).$$

Nella moltiplicazione dei numeri complessi i moduli si moltiplicano e gli argomenti si sommano; quindi

$$z^n = r^n(\cos n\phi + i\sin n\phi),$$

e pertanto

$$\rho = r^n,$$
$$\theta = n\phi.$$

La semiretta ϕ = costante del piano z si trasforma nella semiretta $\theta = n\phi$ = costante del piano w. L'angolo α fra due semirette diventa quindi $\beta = n\alpha$, considerando gli angoli con la loro variazione continua. La trasformazione non è biunivoca. Infatti, ogni punto non nullo w, di modulo ρ e argomento θ, è immagine di n punti con modulo $r = \sqrt[n]{\rho}$ e argomenti

$$\phi = \frac{\theta}{n}, \frac{\theta}{n} + \frac{2\pi}{n}, \cdots, \frac{\theta}{n} + \frac{2\pi}{n}(n-1).$$

Elevando alla potenza n, i moduli diventano tutti ρ e gli argomenti diventano

$$\theta, \theta + 2\pi, \cdots, \theta + 2\pi(n-1),$$

Poiché un multiplo di 2π non modifica la posizione del punto, tutte le immagini coincidono nel piano w.

Trasformazioni conformi.

Se una funzione analitica $w = f(z)$ trasforma biunivocamente un dominio D del piano z in un dominio Δ del piano w, diciamo che realizza una trasformazione conforme di D su Δ.

Il ruolo delle trasformazioni conformi nella teoria delle funzioni e nelle applicazioni dipende anche dal teorema quasi immediato seguente.

Se $\zeta = F(w)$ è analitica in Δ, la funzione composta $F[f(z)]$ è analitica in D. Il risultato segue dall'uguaglianza

$$\frac{\Delta\zeta}{\Delta z} = \frac{\Delta\zeta}{\Delta w} \cdot \frac{\Delta w}{\Delta z}.$$

Poiché $\zeta = F(w)$ e $w = f(z)$ sono analitiche, entrambi i fattori a destra hanno limite; perciò il rapporto $\Delta\zeta/\Delta z$ ha in ogni punto di D il limite univoco $d\zeta/dz$. La funzione composta è dunque analitica.

Il teorema consente di ricondurre lo studio delle funzioni analitiche in Δ a quello delle funzioni analitiche in D. Se D ha una struttura geometrica più semplice, lo studio si semplifica.

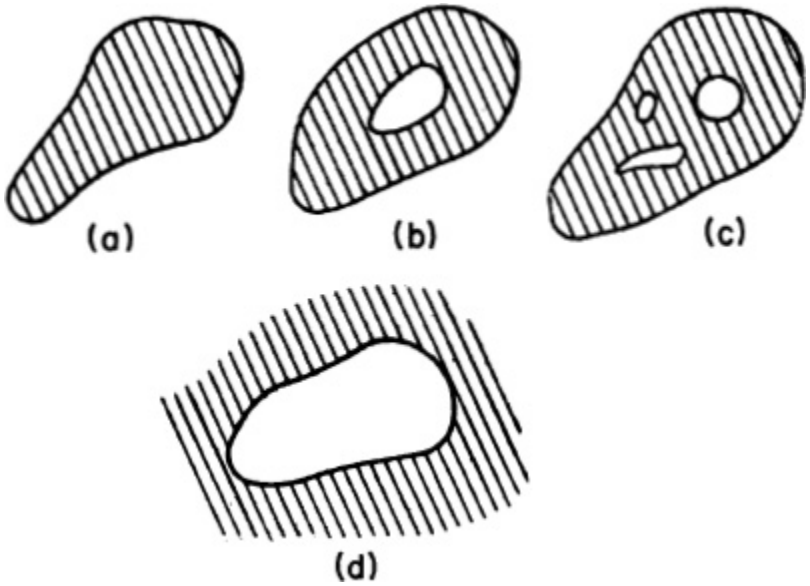


Figura 15.

Una classe particolarmente importante è quella dei domini semplicemente connessi. Nei casi regolari illustrati, il loro contorno ha un solo componente (figura 15a), mentre i domini delle figure 15b e 15c hanno più componenti di contorno. In termini generali, un dominio è semplicemente connesso se ogni curva chiusa in esso può essere contratta a un punto senza uscire dal dominio.

Talvolta interessa studiare funzioni all'esterno di una curva anziché all'interno. Anche il dominio esterno della figura 15d può essere considerato semplicemente connesso, purché vi si includa il punto all'infinito, lavorando sul piano complesso ampliato.

Alla base della teoria delle trasformazioni conformi si trova il seguente notevole teorema di Riemann.

Per ogni dominio semplicemente connesso Δ del piano, diverso dall'intero piano, esiste una funzione analitica che trasforma conformemente il disco unitario in Δ. Si può imporre che il centro vada in un punto assegnato w_0 e che una direzione assegnata nel centro venga trasformata in una direzione assegnata in w_0. Il teorema permette quindi di ricondurre lo studio delle funzioni complesse in tali domini a quello delle funzioni nel disco unitario.

Vediamo in grandi linee come applicare questi fatti alla teoria dell'ala. Supponiamo di voler studiare il flusso intorno a un profilo curvilineo di forma assegnata.

Se costruiamo una trasformazione conforme dell'esterno del profilo sull'esterno del cerchio unitario, possiamo usare la funzione caratteristica del flusso intorno al cerchio per ottenere quella del flusso intorno al profilo.

Indichiamo con ζ il piano del cerchio, con z quello del profilo e con $\zeta = f(z)$ la trasformazione dell'esterno del profilo sull'esterno del cerchio, dove

$$\lim_{z\to\infty} \zeta = \infty.$$

Sia a il punto del cerchio corrispondente al bordo d'uscita A. Costruiamo un flusso con circolazione intorno al cerchio, scegliendo in a il punto dal quale si stacca la linea di corrente (figura 16). Indichiamo la funzione caratteristica con $W(\zeta)$:

$$W(\zeta) = \Phi + i\Psi.$$

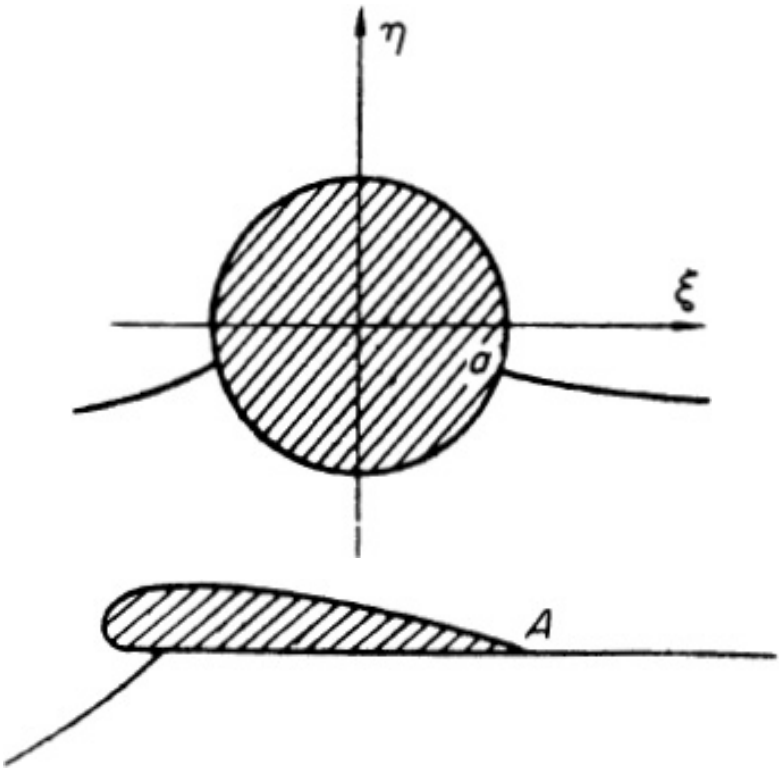


Figura 16.

Le linee di corrente di questo flusso sono date da

$$\Psi = \text{const.}$$

Consideriamo ora la funzione

$$w(z) = W[f(z)],$$

e poniamo

$$w = \phi + i\psi.$$

Mostriamo che $w(z)$ è la funzione caratteristica del flusso intorno al profilo con uscita della corrente dal punto A. Innanzitutto, dobbiamo verificare che il contorno del profilo sia una linea di corrente, cioè che su di esso

$$\psi(x, y) = \text{const.}$$

Ma ciò segue dall'identità

$$\psi(x, y) = \Psi(\xi, \eta),$$

e dal fatto che ai punti (x, y) del profilo corrispondono punti (ξ, η) della circonferenza, sulla quale $\psi(\xi, \eta)$ è costante.

Il punto A corrisponde al punto di ristagno scelto sul cerchio e alla linea di corrente uscente; al bordo affilato la velocità trasformata va valutata mediante il limite appropriato. Scegliendo opportunamente la velocità del flusso intorno al cerchio, si può ottenere un flusso intorno al profilo con la velocità desiderata all'infinito.

L'importanza delle trasformazioni conformi ha suscitato numerosi problemi di costruzione di trasformazioni fra domini di forma geometrica assegnata. In casi semplici ma utili bastano le funzioni elementari; in generale, però, non sono sufficienti. Come abbiamo ricordato, Riemann formulò il teorema generale senza darne una dimostrazione completa secondo i criteri successivi. Una dimostrazione rigorosa richiese il lavoro di molti grandi matematici per diversi decenni.

Dai diversi tentativi di dimostrare il teorema nacquero anche metodi approssimati per costruire trasformazioni conformi. La costruzione esplicita può essere molto difficile. Per studiare proprietà generali delle funzioni, tuttavia, spesso non serve conoscere la trasformazione: bastano alcune sue proprietà geometriche. Da qui è nato un vasto studio della geometria delle trasformazioni conformi. Formuliamo un teorema che ne illustra il carattere.

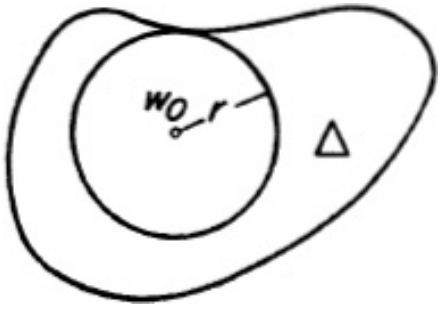


Figura 17.

Il disco unitario del piano z sia trasformato in un dominio Δ (figura 17). Per una trasformazione del tutto arbitraria non possiamo imporre restrizioni analoghe al comportamento nell'origine. Per una trasformazione conforme vale invece il risultato seguente.

Il fattore di dilatazione nell'origine non supera quattro volte il raggio del massimo disco con centro in w_0 contenuto nel dominio:

$$|f'(0)| \leqslant 4r.$$

Numerosi lavori di matematici sovietici hanno studiato questi problemi, ottenendo formule esatte per classi interessanti di trasformazioni, metodi di calcolo approssimato e teoremi geometrici generali.

Trasformazioni quasiconformi.

Le trasformazioni conformi sono strettamente legate alle funzioni analitiche, cioè alle coppie di funzioni che soddisfano le equazioni di Cauchy-Riemann:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y},$$

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Molti problemi della fisica matematica coinvolgono però sistemi differenziali più generali, anch'essi interpretabili come trasformazioni fra piani con specifiche proprietà geometriche locali. Consideriamo, per esempio, il sistema

$$\begin{aligned} \frac{\partial u}{\partial x} &= p(x, y)\frac{\partial v}{\partial y}, \\ \frac{\partial v}{\partial x} &= -p(x, y)\frac{\partial u}{\partial y}. \end{aligned} \qquad (38)$$

Se $p(x, y) = 1$, ritroviamo le equazioni di Cauchy-Riemann. Per una funzione positiva $p(x, y)$, possiamo interpretare ogni soluzione della (38) come una trasformazione dal piano Oxy al

piano *Ouv*. Esaminiamola vicino a (x_0, y_0): mantenendo soltanto i primi termini degli sviluppi in $x - x_0$ e $y - y_0$, otteniamo la trasformazione affine

$$\begin{aligned} u - u_0 &= \left(\frac{\partial u}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial u}{\partial y}\right)_0 (y - y_0), \\ v - v_0 &= \left(\frac{\partial v}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial v}{\partial y}\right)_0 (y - y_0). \end{aligned} \tag{39}$$

Se u e v soddisfano la (38), questa trasformazione, nei punti non degeneri, ha la proprietà seguente.

Le ellissi centrate in (x_0, y_0), con assi paralleli agli assi coordinati e rapporto dei semiassi

$$\frac{b}{a} = p(x_0 , y_0)$$

vengono trasformate in circonferenze centrate in (u_0, v_0) nel piano *Ouv*.

Dimostriamolo. L'equazione di una circonferenza con questo centro è

$$(u - u_0)^2 + (v - v_0)^2 = \rho^2.$$

Sostituendo le espressioni di $u - u_0$ e $v - v_0$, otteniamo l'equazione della curva corrispondente nel piano *Oxy*:

$$\begin{aligned} &\left[\left(\frac{\partial u}{\partial x}\right)_0^2 + \left(\frac{\partial v}{\partial x}\right)_0^2\right](x - x_0)^2 \\ &\quad + 2\left[\left(\frac{\partial u}{\partial x}\right)_0\left(\frac{\partial u}{\partial y}\right)_0 + \left(\frac{\partial v}{\partial x}\right)_0\left(\frac{\partial v}{\partial y}\right)_0\right](x - x_0)(y - y_0) \\ &\quad + \left[\left(\frac{\partial u}{\partial y}\right)_0^2 + \left(\frac{\partial v}{\partial y}\right)_0^2\right](y - y_0)^2 = \rho^2. \end{aligned}$$

Usiamo le equazioni (38) per esprimere le derivate di v mediante quelle di u:

$$\left[\left(\frac{\partial u}{\partial x}\right)_0^2 + p^2\left(\frac{\partial u}{\partial y}\right)_0^2\right](x - x_0)^2 + \frac{1}{p^2}\left[\left(\frac{\partial u}{\partial x}\right)_0^2 + p^2\left(\frac{\partial u}{\partial y}\right)_0^2\right](y - y_0)^2 = \rho^2.$$

Ponendo

$$\begin{aligned} a &= \frac{\rho}{\sqrt{\left(\frac{\partial u}{\partial x}\right)_0^2 + p^2\left(\frac{\partial u}{\partial y}\right)_0^2}}, \\ b &= \frac{p\rho}{\sqrt{\left(\frac{\partial u}{\partial x}\right)_0^2 + p^2\left(\frac{\partial u}{\partial y}\right)_0^2}}, \end{aligned}$$

l'equazione diventa

$$\frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} = 1.$$

La curva trasformata in una circonferenza è dunque un'ellisse con le proprietà indicate.

Se consideriamo la trasformazione esatta anziché l'approssimazione affine, la proprietà diventa tanto più precisa quanto più piccole sono le ellissi. Si può quindi dire che essa vale per ellissi infinitesime.

Le equazioni (38) determinano, in ogni punto non degenere, l'orientazione e il rapporto dei semiassi dell'ellisse infinitesima trasformata in una circonferenza. Con le opportune condizioni di orientazione, questa proprietà geometrica caratterizza a sua volta il sistema differenziale. Lo studio delle sue soluzioni può così essere tradotto nello studio di trasformazioni con proprietà geometriche assegnate.

Per le equazioni di Cauchy-Riemann la proprietà si esprime in particolare così:

Una circonferenza infinitesima centrata in (x_0, y_0) diventa una circonferenza infinitesima centrata in (u_0, v_0), nei punti in cui la derivata non si annulla.

Una classe assai ampia di equazioni della fisica matematica può essere ricondotta allo studio di trasformazioni con le proprietà geometriche seguenti.

In ogni punto (x, y) del piano dell'argomento si assegnano due ellissi, specificando la direzione dei loro semiassi e il rapporto delle lunghezze. Si cerca una trasformazione dal piano *Oxy* al piano *Ouv* che trasformi un'ellisse infinitesima della prima famiglia in una della seconda, centrata nel punto corrispondente (u, v).

Trasformazioni legate a sistemi più generali di questo genere furono introdotte dal matematico sovietico M. A. Lavrent'ev e chiamate quasiconformi. Lo studio geometrico di trasformazioni definite da sistemi differenziali permise di estendere i metodi delle funzioni analitiche a una classe molto vasta di problemi. Lavrent'ev e i suoi allievi svilupparono questa teoria e ne trovarono numerose applicazioni alla fisica matematica, alla meccanica e alla geometria. È interessante che le trasformazioni quasiconformi si siano rivelate feconde anche nello studio delle stesse funzioni analitiche. Non possiamo qui soffermarci sulle molte applicazioni del metodo geometrico.

§4. L'integrale curvilineo; la formula di Cauchy e le sue conseguenze

Integrali di funzioni complesse.

Nello studio delle funzioni analitiche ha un ruolo fondamentale l'integrazione complessa. All'integrale definito di una funzione reale corrisponde l'integrale di una funzione complessa lungo una curva. Consideriamo una curva C che va da z_0 a z e una funzione f definita in un dominio contenente C. Suddividiamo la curva in piccoli archi mediante i punti (figura 18)

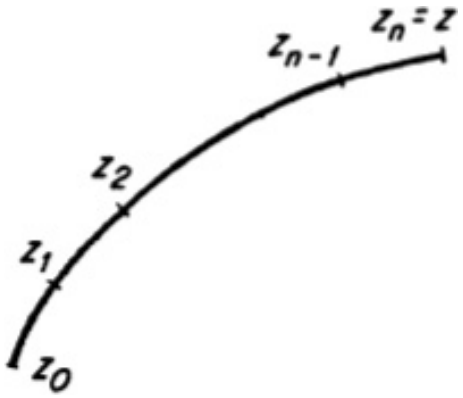


Figura 18.

$$z_0 , z_1 , \cdots, z_n = z$$

e consideriamo la somma

$$S = \sum_{k=1}^{n} f(z_k)\,(z_k - z_{k-1}).$$

Se f è continua e C ha lunghezza finita, si dimostra, come nel caso reale, che aumentando il numero dei punti e facendo tendere a zero la lunghezza massima degli archi della suddivisione, la somma S tende a un limite determinato. Esso si chiama integrale lungo la curva C e si indica con

$$\int_C f(z)\,dz.$$

Nella definizione abbiamo distinto l'inizio dalla fine della curva: abbiamo cioè scelto un verso di percorrenza, o orientazione.

Si dimostrano facilmente alcune semplici proprietà dell'integrale.

1. L'integrale della somma di due funzioni è uguale alla somma dei loro integrali:

$$\int_C [f(z) + g(z)]\,dz = \int_C f(z)\,dz + \int_C g(z)\,dz.$$

2. Un fattore costante può essere portato fuori dal segno d'integrale:

$$\int_C Af(z)\,dz = A\int_C f(z)\,dz.$$

3. Se C è ottenuta concatenando le curve C_1 e C_2, allora

$$\int_C f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz.$$

4. Se $\bar{C}$ indica la curva C percorsa nel verso opposto, allora

$$\int_{\bar{C}} f(z)\,dz = -\int_C f(z)\,dz.$$

Queste proprietà sono immediate per le somme approssimanti e si trasferiscono all'integrale passando al limite.

5. Se la curva C ha lunghezza L e su di essa vale ovunque

$$|f(z)| \leqslant M$$

allora

$$\left|\int_C f(z)\,dz\right| \leqslant ML.$$

Dimostriamo quest'ultima proprietà. Basta provarla per la somma S, perché l'ineguaglianza si conserverà al limite. Per la somma si ha

$$|S| = \left|\sum f(z_k)(z_k - z_{k-1})\right| \leqslant \sum |f(z_k)|\,|z_k - z_{k-1}| \leqslant M\sum |z_k - z_{k-1}|.$$

La somma che compare nel secondo fattore è la lunghezza della spezzata inscritta in C, con vertici nei punti z_k. Questa lunghezza non supera quella della curva; dunque

$$|S| \leqslant ML.$$

Consideriamo l'integrale della funzione più semplice, $f(z) = 1$. Evidentemente,

$$S = (z_1 - z_0) + (z_2 - z_1) + \cdots + (z_n - z_{n-1}) = z_n - z_0 = z - z_0.$$

e quindi

$$\int_C 1 \cdot dz = z - z_0.$$

Per $f(z) = 1$ l'integrale ha lo stesso valore su tutte le curve che congiungono z_0 e z: dipende soltanto dagli estremi del cammino. È però facile vedere che ciò non vale per una funzione complessa arbitraria. Per esempio, se $f(z) = x$, un semplice calcolo dà

$$\int_{C_1} x\,dz = \frac{x^2}{2} + iyx, \quad \int_{C_2} x\,dz = \frac{x^2}{2}, \quad z = x + iy,$$

dove C_1 e C_2 sono i cammini della figura 19.

Lasciamo al lettore la verifica di queste uguaglianze.

Un fatto fondamentale della teoria delle funzioni analitiche è espresso dal teorema di Cauchy.

Se $f(z)$ è derivabile in ogni punto di un dominio semplicemente connesso D del piano, gli integrali lungo tutti i cammini ammissibili contenuti in D che congiungono due punti assegnati z_0 e z sono uguali.

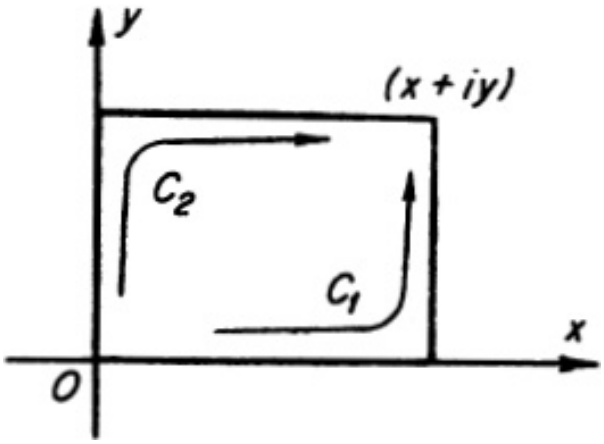


Figura 19.

Non dimostreremo qui il teorema; il lettore interessato potrà consultare un corso di funzioni di variabile complessa. Ne ricordiamo alcune conseguenze importanti.

Il teorema permette innanzitutto di introdurre l'integrale indefinito di una funzione analitica. Fissato z_0, consideriamo l'integrale lungo una curva che lo congiunge a z:

$$F(z) = \int_{z_0}^{z} f(\zeta)\,d\zeta.$$

Possiamo scegliere qualsiasi cammino contenuto nel dominio, poiché il suo cambiamento non modifica l'integrale, che dipende soltanto da z. La funzione $F(z)$ così ottenuta si chiama integrale indefinito, o primitiva, di $f(z)$.

La derivata di una primitiva di $f(z)$ è uguale a $f(z)$.

Per molte applicazioni è comoda una formulazione leggermente diversa del teorema di Cauchy:

Se $f(z)$ è derivabile ovunque in un dominio semplicemente connesso, il suo integrale lungo ogni contorno chiuso contenuto nel dominio è nullo:

$$\int_\Gamma f(z)\,dz = 0.$$

Infatti, gli estremi di un cammino chiuso coincidono, e gli stessi estremi possono essere congiunti dal cammino costante.

Per i contorni semplici chiusi assumeremo come orientazione positiva quella antioraria. Il contorno percorso in senso orario sarà indicato con $\bar{\Gamma}$.

L'integrale di Cauchy.

Dal teorema precedente si deduce la formula fondamentale di Cauchy, che esprime i valori interni di una funzione analitica mediante i suoi valori sul contorno. Per un contorno semplice chiuso positivamente orientato, con la funzione analitica anche in un intorno del contorno e del suo interno, vale

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)\,d\zeta}{\zeta - z}.$$

Dimostriamo la formula. Fissiamo z e consideriamo ζ come variabile indipendente. La funzione

$$\phi(\zeta) = \frac{f(\zeta)}{\zeta - z}$$

è continua e derivabile in ogni punto ζ del dominio, salvo $\zeta = z$, dove il denominatore si annulla. Questa eccezione impedisce di applicare direttamente il teorema di Cauchy alla funzione $\phi(\zeta)$ sul contorno C.

Prendiamo una piccola circonferenza K_ρ di centro z e raggio ρ e mostriamo che

$$\int_C \phi(\zeta)\,d\zeta = \int_{K_\rho} \phi(\zeta)\,d\zeta. \qquad (40)$$

Costruiamo a questo scopo il contorno ausiliario Γ_ρ, composto da C, da un cammino γ_ρ che lo collega alla circonferenza, dalla circonferenza $\bar{K}_\rho$ percorsa nel verso opposto e dal cammino di ritorno (figura 20). Le frecce indicano l'orientazione. Il punto $\zeta = z$ è escluso dall'interno del contorno tagliato; la funzione è dunque analitica nella regione considerata e

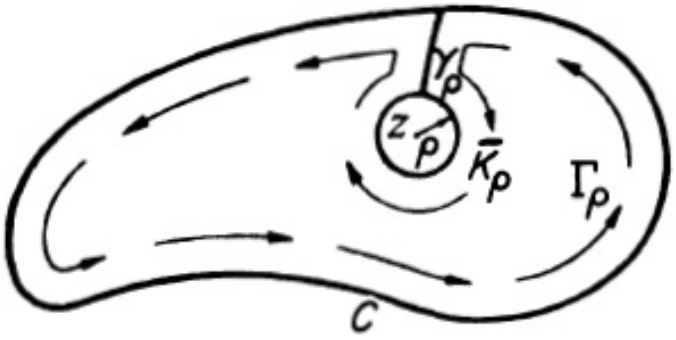


Figura 20.

$$\int_{\Gamma_\rho} \phi(\zeta)\, d\zeta = 0. \qquad (41)$$

Il contorno è formato dai quattro tratti C, γ_ρ, $\bar{K}_\rho$ e $\bar{\gamma}_\rho$. Per la proprietà 3,

$$\int_{\Gamma_\rho} \phi(\zeta)\, d\zeta = \int_C \phi(\zeta)\, d\zeta + \int_{\gamma_\rho} \phi(\zeta)\, d\zeta + \int_{\bar{K}_\rho} \phi(\zeta)\, d\zeta + \int_{\bar{\gamma}_\rho} \phi(\zeta)\, d\zeta = 0.$$

Sostituendo agli integrali lungo $\bar{K}_\rho$ e $\bar{\gamma}_\rho$ quelli lungo K_ρ e γ_ρ, con il segno cambiato secondo la proprietà 4, otteniamo

$$\int_{\Gamma_\rho} \phi(\zeta)\, d\zeta = \int_C \phi(\zeta)\, d\zeta - \int_{K_\rho} \phi(\zeta)\, d\zeta = 0,$$

e la (40) è dimostrata.

Per calcolare il secondo membro della (40), scriviamo

$$\int_{K_\rho} \phi(\zeta)\, d\zeta = \int_{K_\rho} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \int_{K_\rho} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta + \int_{K_\rho} \frac{f(z)\, d\zeta}{\zeta - z}$$

$$= \int_{K_\rho} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta + f(z) \int_{K_\rho} \frac{d\zeta}{\zeta - z}. \qquad (42)$$

Calcoliamo prima il secondo termine. Sulla circonferenza K_ρ,

$$\zeta = z + \rho(\cos\theta + i \sin\theta).$$

Poiché z e ρ sono costanti,

$$d\zeta = \rho(-\sin\theta + i\cos\theta)\, d\theta = i\rho(\cos\theta + i\sin\theta)\, d\theta,$$

e quindi

$$\zeta - z = \rho(\cos\theta + i\sin\theta),$$

da cui

$$\int_{K_\rho} \frac{d\zeta}{\zeta - z} = \int_{K_\rho} i\, d\theta = 2\pi i,$$

poiché durante un giro l'angolo θ aumenta complessivamente di 2π. Dalle (40) e (42) segue

$$\int_C \frac{f(\zeta)\, d\zeta}{\zeta - z} = 2\pi i f(z) + \int_{K_\rho} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta.$$

Facciamo tendere ρ a zero. Il primo membro e il primo termine del secondo membro non cambiano. Mostriamo che l'altro termine tende a zero: otterremo così la formula di Cauchy. Osserviamo che

$$\lim_{\zeta \to z} [f(\zeta) - f(z)]/(\zeta - z) = f'(z),$$

L'espressione integranda ha quindi limite finito ed è limitata:

$$\left| \frac{f(\zeta) - f(z)}{\zeta - z} \right| < M.$$

Applicando la proprietà 5 dell'integrale, otteniamo

$$\left| \int_{K_\rho} \frac{f(\zeta) - f(z)}{\zeta - z}\, d\zeta \right| \leqslant M 2\pi\rho \to 0.$$

La formula di Cauchy è così dimostrata. Essa è uno degli strumenti fondamentali della teoria delle funzioni complesse.

Sviluppo in serie di potenze delle funzioni derivabili.

Applichiamo la formula di Cauchy per stabilire due proprietà fondamentali delle funzioni complesse derivabili.

Una funzione complessa che possiede la derivata prima in ogni punto di un dominio D possiede in esso derivate di tutti gli ordini.

Infatti, all'interno di un opportuno contorno chiuso la funzione si esprime mediante l'integrale di Cauchy:

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{\zeta - z}\, d\zeta.$$

La funzione di z sotto il segno d'integrale è derivabile; derivando sotto il segno, si ha

$$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)}{(\zeta - z)^2}\, d\zeta.$$

L'integranda è ancora derivabile, e una nuova derivazione dà

$$f''(z) = \frac{1 \cdot 2}{2\pi i} \int_C \frac{f(\zeta)\, d\zeta}{(\zeta - z)^3}.$$

Proseguendo, si ottiene la formula generale

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(\zeta)\, d\zeta}{(\zeta - z)^{n+1}}.$$

Possiamo così calcolare la derivata di qualsiasi ordine. Per rendere la dimostrazione completamente rigorosa occorre giustificare la derivazione sotto il segno d'integrale; omettiamo questa parte.

La seconda proprietà è la seguente.

Se $f(z)$ è derivabile in tutti i punti di un disco K di centro a, allora si può sviluppare nella serie di Taylor

$$f(z) = f(a) + \frac{f'(a)}{1!}(z - a) + \cdots + \frac{f^{(n)}(a)}{n!}(z - a)^n + \cdots.$$

convergente in tutto l'interno di K.

Nel §1 abbiamo introdotto le funzioni analitiche come funzioni sviluppabili in serie di potenze. Il teorema afferma che ogni funzione complessa derivabile in un dominio è analitica. È una proprietà peculiare del campo complesso: nel campo reale, una funzione che possiede la derivata prima può non avere la derivata seconda.

Dimostriamo il teorema appena formulato.

Supponiamo dapprima che f sia analitica all'interno e in un intorno del bordo del disco K, di centro a. Nel suo interno si può scrivere

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\zeta)\, d\zeta}{\zeta - z}. \tag{43}$$

Poniamo

$$\zeta - z = (\zeta - a) - (z - a),$$

allora

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - a) - (z - a)} = \frac{1}{\zeta - a} \frac{1}{1 - \dfrac{z - a}{\zeta - a}}. \tag{44}$$

Poiché z è interno al disco e la variabile d'integrazione ζ si trova sulla circonferenza,

$$\left|\frac{z-a}{\zeta-a}\right| < 1,$$

e dalla formula della serie geometrica segue

$$\frac{1}{1-\dfrac{z-a}{\zeta-a}} = 1 + \left(\frac{z-a}{\zeta-a}\right) + \cdots + \left(\frac{z-a}{\zeta-a}\right)^n + \cdots, \qquad (45)$$

con serie convergente a destra. Usando le (44) e (45), la (43) diventa

$$f(z) = \frac{1}{2\pi i}\int_C \left[\frac{f(\zeta)}{\zeta-a} + (z-a)\frac{f(\zeta)}{(\zeta-a)^2} + \cdots + (z-a)^n \frac{f(\zeta)}{(\zeta-a)^{n+1}} + \cdots\right] d\zeta.$$

Integriamo ora la serie termine per termine, operazione che può essere giustificata rigorosamente. In ogni termine portiamo fuori dall'integrale il fattore $(z - a)^n$, indipendente da ζ:

$$f(z) = \frac{1}{2\pi i}\int_C \frac{f(\zeta)\,d\zeta}{\zeta-a} + \frac{z-a}{2\pi i}\int_C \frac{f(\zeta)\,d\zeta}{(\zeta-a)^2} + \cdots + \frac{(z-a)^n}{2\pi i}\int_C \frac{f(\zeta)\,d\zeta}{(\zeta-a)^{n+1}} + \cdots.$$

Usando le formule integrali delle derivate successive, possiamo scrivere

$$\frac{1}{2\pi i}\int_C \frac{f(\zeta)\,d\zeta}{(\zeta-a)^{n+1}} = \frac{f^{(n)}(a)}{n!},$$

e otteniamo infine

$$f(z) = f(a) + \frac{f'(a)}{1!}(z-a) + \cdots + \frac{f^{(n)}(a)}{n!}(z-a)^n + \cdots.$$

Abbiamo dunque mostrato che le funzioni complesse derivabili sono sviluppabili in serie di potenze. Viceversa, le funzioni rappresentate da serie di potenze sono derivabili all'interno del disco di convergenza, e le derivate si ottengono derivando termine per termine. Anche questa operazione può essere giustificata rigorosamente. Per una funzione analitica soltanto nell'interno di K, la dimostrazione si applica a circonferenze concentriche più piccole.

Funzioni intere.

Una serie di potenze rappresenta analiticamente una funzione in un disco. Se il raggio è finito, esso è uguale alla distanza del centro dalla singolarità più vicina, cioè dal punto più vicino in cui la funzione cessa di essere analitica.

È naturale distinguere le funzioni analitiche per tutti i valori finiti dell'argomento. Esse sono rappresentate da serie di potenze convergenti per ogni z e si chiamano funzioni intere. Sviluppando intorno all'origine, una funzione intera si esprime nella forma

$$G(z) = c_0 + c_1 z + c_2 z^2 + \cdots + c_n z^n + \cdots.$$

Se da un certo indice in poi tutti i coefficienti sono nulli, la funzione è semplicemente un polinomio, detto anche funzione razionale intera:

$$P(z) = c_0 + c_1 z + \cdots + c_n z^n.$$

Se invece vi sono infiniti termini non nulli, la funzione intera si chiama trascendente.

Ne sono esempi

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots,$$

$$\sin z = \frac{z}{1!} - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots,$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \cdots.$$

Nello studio dei polinomi ha grande importanza la distribuzione delle radici dell'equazione

$$P(z) = 0,$$

o, più generalmente, dei punti nei quali il polinomio assume un valore assegnato A:

$$P(z) = A.$$

Il teorema fondamentale dell'algebra afferma che ogni polinomio non costante assume ogni valore A almeno una volta. Ciò non si estende a tutte le funzioni intere: per esempio, e^z non assume mai il valore zero. Vale però il teorema di Picard nella seguente forma per le funzioni intere trascendenti: ogni valore assegnato è assunto infinite volte, con al più un valore eccezionale.

La distribuzione dei punti nei quali una funzione intera assume un valore A è uno dei problemi centrali della teoria delle funzioni intere.

Il numero di radici di un polinomio, contate con molteplicità, coincide con il grado. Il grado è strettamente legato alla rapidità di crescita di $|P(z)|$ per $|z| \to \infty$. Possiamo infatti scrivere

$$|P(z)| = |z|^n \cdot \left| a_n + \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n} \right|,$$

Poiché il secondo fattore tende a $|a_n|$, per grandi $|z|$ un polinomio di grado n cresce come $|a_n|$ $|z|^n$. A un grado più elevato corrispondono quindi crescita più rapida e più radici. Legami analoghi esistono per le funzioni intere. Per una funzione intera trascendente le radici sono in

generale infinite, e contarle tutte non dà un numero finito. Possiamo però contare le radici $n(r, a)$ dell'equazione

$$f(z) = a$$

contenute nel disco di raggio r e studiare come il loro numero cresce con r. Questa crescita è collegata a quella di $M(r)$, massimo del modulo della funzione sulla circonferenza di raggio r. Come abbiamo detto, può esistere un valore eccezionale a per il quale non vi sono radici. La teoria precisa, sotto le proprie ipotesi, i rapporti fra la distribuzione dei valori e la crescita di $\ln M(r)$; non possiamo formulare qui queste leggi in modo completo.

Le proprietà della distribuzione degli zeri delle funzioni intere sono collegate alla teoria dei numeri. Esse hanno consentito di stabilire importanti proprietà della funzione zeta di Riemann,[6] sulle quali si fondano numerosi teoremi sui numeri primi.

Funzioni fratte o meromorfe.

La classe delle funzioni intere estende quella dei polinomi. Dai polinomi si passa alla classe più ampia delle funzioni razionali,

$$R(z) = \frac{P(z)}{Q(z)},$$

che sono quozienti di due polinomi.

È naturale costruire analogamente una nuova classe a partire dalle funzioni intere. Una funzione $f(z)$ che sia quoziente di due funzioni intere G_1 e G_2, con denominatore non identicamente nullo,

$$f(z) = \frac{G_1(z)}{G_2(z)}$$

si chiama funzione fratta o meromorfa nel piano. Questa classe ha un ruolo importante nell'analisi. Fra le funzioni elementari meromorfe troviamo, per esempio,

$$\tan z = \frac{\sin z}{\cos z}, \qquad \cot z = \frac{\cos z}{\sin z}.$$

Una funzione meromorfa non è necessariamente analitica in tutto il piano. Nei punti in cui il denominatore si annulla, dopo le eventuali cancellazioni con il numeratore, possono comparire poli. Gli zeri del denominatore sono isolati. Vicino a un polo a la funzione non può essere sviluppata in serie di Taylor; ammette però uno sviluppo contenente anche un numero finito di potenze negative di $z - a$:

$$f(z) = \frac{C_{-m}}{(z-a)^m} + \cdots + \frac{C_{-1}}{z-a} + C_0 + C_1(z-a) + \cdots + C_n(z-a)^n + \cdots. \tag{46}$$

Avvicinandosi al polo a, il modulo di $f(z)$ tende all'infinito. Un punto singolare isolato con questa proprietà si chiama polo. La perdita di analiticità è dovuta ai termini con potenze negative nello sviluppo (46). L'espressione

$$\frac{C_{-m}}{(z-a)^m} + \cdots + \frac{C_{-1}}{(z-a)}$$

descrive la parte singolare del comportamento vicino al polo e si chiama parte principale dello sviluppo. Conoscendo le parti principali in tutti i poli, si può ricostruire una funzione meromorfa a meno di una funzione intera, con gli opportuni accorgimenti di convergenza. Per esempio, se f è razionale e tende a zero all'infinito, è uguale alla somma delle parti principali nei suoi poli, che sono in numero finito:

$$f(z) = \sum_{(k)} \left[\frac{C_{-m_k}^{(k)}}{(z-a_k)^{m_k}} + \cdots + \frac{C_{-1}^{(k)}}{z-a_k} \right].$$

In generale, una funzione razionale è la somma delle parti principali e di un polinomio:

$$f(z) = \sum_{(k)} \left[\frac{C_{-m_k}^{(k)}}{(z-a_k)^{m_k}} + \cdots + \frac{C_{-1}^{(k)}}{z-a_k} \right] + C_0 + C_1 z + \cdots + C_m z^m. \tag{47}$$

La formula (47) rende evidente il ruolo delle singolarità. È comoda nelle applicazioni e importante teoricamente, perché mostra come i punti singolari contribuiscono alla struttura della funzione. Una costruzione analoga vale per le funzioni meromorfe, includendo gli eventuali termini correttivi. Riportiamo senza dimostrazione l'esempio di cot z. I suoi poli sono gli zeri di

$$\sin z = 0$$

e si trovano nei punti ..., $-k\pi$, ..., $-\pi$, 0, π, ..., $k\pi$, ... Si dimostra che la parte principale nel polo $z = k\pi$ è

$$\frac{1}{z-k\pi},$$

e che cot z si esprime mediante la somma delle parti principali, raggruppate simmetricamente come nella formula:

$$\cot z = \frac{1}{z} + \sum_{k=1}^{\infty} \left(\frac{1}{z-k\pi} + \frac{1}{z+k\pi} \right). \tag{48}$$

Lo sviluppo mediante parti principali mostra chiaramente la posizione delle singolarità e, con le opportune condizioni di convergenza, consente di calcolare la funzione in tutto il suo dominio.

La teoria delle funzioni meromorfe è fondamentale per molte classi di funzioni importanti nell'analisi. Va sottolineato il suo significato per le equazioni della fisica matematica. La teoria delle equazioni integrali, che risponde a numerose questioni fondamentali in questo campo, si sviluppò in larga parte grazie ai teoremi sulle funzioni meromorfe.

In seguito, anche lo sviluppo della teoria degli operatori, la parte dell'analisi funzionale più strettamente connessa con la fisica matematica, si è spesso basato su risultati della teoria delle funzioni analitiche.

La rappresentazione analitica delle funzioni.

Abbiamo visto che, vicino a ogni punto di analiticità, una funzione è rappresentata da una serie di potenze. Per una funzione intera la serie converge in tutto il piano e fornisce un'espressione valida ovunque. Se la funzione non è intera, la serie di Taylor converge soltanto fino alla singolarità più vicina; non permette quindi di calcolarla in tutto il dominio. Per una funzione meromorfa, uno sviluppo mediante parti principali, completato quando necessario da una funzione intera e da termini di convergenza, può invece fornire una rappresentazione globale.

Se una funzione è definita in un disco, oppure vogliamo studiarla soltanto in un disco contenuto nel suo dominio, la serie di Taylor è adatta. Per domini di forma diversa nasce però il problema di trovare un'espressione valida in tutto il dominio. I termini di una serie di potenze sono i polinomi più semplici, $a_n z^n$. È naturale chiedersi se, in un dominio più generale, si possa sviluppare la funzione in una serie di polinomi più generali. Ogni termine sarebbe ancora calcolabile mediante sole operazioni aritmetiche, offrendo una rappresentazione fondata sulle operazioni più elementari. Una risposta generale è data dal teorema seguente.

Una funzione analitica in un dominio semplicemente connesso del piano può essere sviluppata in una serie di polinomi, convergente uniformemente su ogni compatto del dominio:

$$f(z) = P_0(z) + P_1(z) + \cdots + P_n(z) + \cdots.$$

Il teorema dà una risposta generale, ma non fornisce ancora una costruzione esplicita della serie per una funzione assegnata, come avviene per Taylor. In questo senso pone il problema più che risolverlo. Costruire la serie a partire dalla funzione o da alcune sue proprietà, ottenere una convergenza più rapida, collegare lo sviluppo al comportamento della funzione e studiare la struttura di una funzione definita da una serie assegnata sono tutti problemi della vasta teoria dell'approssimazione mediante polinomi. Alla sua costruzione contribuirono in misura importante i matematici sovietici, con numerosi risultati fondamentali.

§5. Unicità e prolungamento analitico

Proprietà di unicità delle funzioni analitiche.

Una delle proprietà più notevoli è espressa dal teorema seguente.

Se due funzioni analitiche in un dominio connesso *D* coincidono su un arco di curva contenuto nel dominio, esse coincidono in tutto *D*.

La dimostrazione è molto semplice. Siano $f_1(z)$ e $f_2(z)$ analitiche in *D* e coincidenti sulla curva *C*. La differenza

$$\phi(z) = f_1(z) - f_2(z)$$

è analitica in *D* e nulla su *C*. Mostriamo che $\phi(z) = 0$ in tutto il dominio. Se esistesse un punto z_0 in cui $\phi(z_0) \neq 0$, prolungheremmo la curva fino a esso, rimanendo in *D* (figura 21). Percorriamo la curva prolungata Γ finché la funzione rimane nulla, e sia ζ l'ultimo punto raggiungibile in questo modo. Per continuità, $\zeta \neq z_0$; non possiamo proseguire oltre ζ mantenendo l'annullamento su un intero tratto. Mostriamo che ciò è impossibile. Nel tratto Γ_ζ precedente ζ la funzione è identicamente nulla. Le sue derivate possono essere calcolate usando soltanto valori su quel tratto; anch'esse sono quindi nulle, e per continuità lo sono in ζ. In particolare,

$$\phi(\zeta) = \phi'(\zeta) = \cdots = \phi^{(n)}(\zeta) = \cdots = 0.$$

Sviluppiamo $\phi(z)$ in serie di Taylor intorno a ζ. Tutti i coefficienti sono nulli, perciò

$$\phi(z) = 0$$

in un disco di centro ζ contenuto in *D*. La funzione è dunque nulla anche su un tratto di Γ oltre ζ, contraddicendo la sua definizione. L'ipotesi $\phi(z_0) \neq 0$ è impossibile.

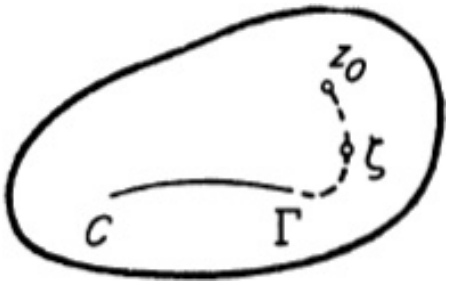


Figura 21.

Conoscendo i valori di una funzione analitica su un arco o su una parte aperta del dominio, i suoi valori sono quindi determinati univocamente in tutto il dominio connesso. I valori nelle diverse parti del piano sono strettamente collegati fra loro.

Per comprendere l'importanza di questa proprietà basta ricordare che la definizione generale di funzione complessa ammette qualsiasi legge di corrispondenza fra argomenti e valori. In generale non è possibile determinare i valori in una regione conoscendoli altrove. La sola richiesta di derivabilità complessa è dunque tanto forte da imporre un legame fra valori in regioni diverse.

Nella teoria delle funzioni reali la derivabilità non ha da sola conseguenze analoghe. Si possono costruire funzioni infinitamente derivabili che coincidono su una parte dell'asse *Ox* e differiscono altrove. Una funzione nulla per $x \leq 0$ può, per esempio, essere diversa da zero per $x > 0$ e avere derivate continue di ogni ordine. Basta porre, per $x > 0$,

$$f(x) = e^{-1/x}.$$

Prolungamento analitico e funzioni analitiche complete.

Il dominio di una funzione complessa è spesso limitato dal modo stesso in cui viene definita. Consideriamo l'esempio elementare della serie

$$f(z) = 1 + z + z^2 + \cdots + z^n + \cdots. \qquad (49)$$

Essa converge nel disco unitario e diverge al suo esterno, oltre che sulla circonferenza. La funzione analitica definita dalla (49) è dunque inizialmente assegnata soltanto nel disco. Sappiamo però che, per $|z| < 1$, la somma è

$$f(z) = \frac{1}{1 - z}. \qquad (50)$$

La formula (50) ha senso per ogni $z \neq 1$. Per il teorema di unicità, essa rappresenta l'unica estensione analitica sul piano privato di 1 che coincide con la somma della (49) nel disco unitario. La funzione, definita inizialmente soltanto nel disco, è così prolungata a tutto il piano tranne il punto 1.

Se $f(z)$ è analitica in un dominio D ed esiste una funzione analitica $F(z)$ in un dominio connesso più grande Δ, contenente D, che coincide con f in D, allora il teorema di unicità determina F in modo univoco in Δ.

La funzione F si chiama prolungamento analitico di f. Una funzione analitica si dice completa quando non può essere ulteriormente prolungata oltre il dominio sul quale è già definita. Per esempio, una funzione intera è completa nel piano; una funzione meromorfa nel piano, considerata analitica fuori dai suoi poli, non è prolungabile analiticamente attraverso questi ultimi. Esistono però anche funzioni il cui dominio completo di analiticità è limitato. Non presenteremo questi esempi più complessi.

Il concetto di funzione analitica completa conduce alla necessità di considerare funzioni complesse a più valori. Vediamolo sull'esempio

$$\operatorname{Ln} z = \ln r + i\phi,$$

dove $r = |z|$ e $\phi = \arg z$. In un punto $z_0 = r_0(\cos \phi_0 + i \sin \phi_0)$ scegliamo il valore iniziale

$$(\mathrm{Ln}\ z)_0 = \ln r_0 + i\phi_0\,,$$

e prolunghiamo la funzione lungo una curva C. Come abbiamo già osservato, se z percorre una curva semplice chiusa C_0 partendo da z_0, compiendo un giro antiorario intorno all'origine e tornando al punto iniziale (figura 22), $\ln r_0$ riprende il valore iniziale, mentre l'argomento aumenta di 2π. Il prolungamento di Ln z lungo il giro ne aumenta quindi il valore di $2\pi i$. Dopo n giri, al posto del valore iniziale

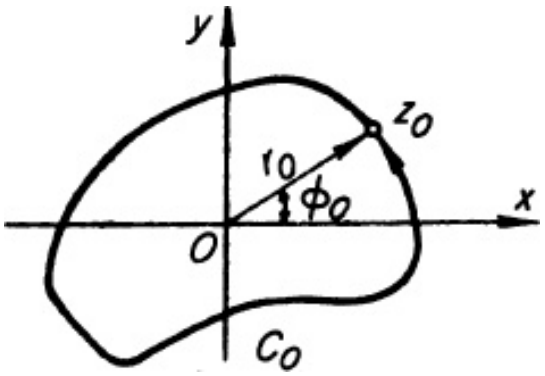


Figura 22.

$$(\mathrm{Ln}\ z)_0 = \ln r_0 + i\phi_0$$

otteniamo

$$(\mathrm{Ln}\ z)_n = \ln r_0 + (2\pi n + \phi_0)i.$$

Dopo m giri nel verso opposto, otteniamo invece

$$(\mathrm{Ln}\ z)_{-m} = \ln r_0 + (-2\pi m + \phi_0)i.$$

Queste osservazioni mostrano la necessità di considerare il legame fra i diversi valori di Ln z, che sono infiniti. Un ruolo particolare spetta all'origine, intorno alla quale si passa da un valore all'altro. Se invece si percorre una curva semplice chiusa che non circonda l'origine, il valore non cambia. Il punto $z = 0$ è un punto di diramazione di Ln z.

In generale, se il prolungamento di una funzione $f(z)$ lungo un giro intorno ad a porta da un suo valore a un altro, a si chiama punto di diramazione.

Consideriamo un secondo esempio. Sia

$$w = \sqrt[n]{z}.$$

Come abbiamo già osservato, anche questa funzione è a più valori: per $z \neq 0$ assume n valori,

$$\sqrt[n]{r}\left(\cos\frac{\phi}{n}+i\sin\frac{\phi}{n}\right),\qquad \sqrt[n]{r}\left(\cos\frac{\phi+2\pi}{n}+i\sin\frac{\phi+2\pi}{n}\right),$$
$$\cdots,\ \sqrt[n]{r}\left(\cos\frac{\phi+2\pi(n-1)}{n}+i\sin\frac{\phi+2\pi(n-1)}{n}\right).$$

Tutti possono essere ottenuti a partire da uno solo,

$$w_0=\sqrt[n]{r_0}\left(\cos\frac{\phi_0}{n}+i\sin\frac{\phi_0}{n}\right)$$

percorrendo curve chiuse intorno all'origine, poiché ogni giro antiorario aumenta l'argomento di 2π.

Dopo $n-1$ giri, a partire dal primo valore di $\sqrt[n]{z}$, otteniamo tutti gli altri $n-1$ valori. Al giro successivo, il numero n, si ottiene

$$\sqrt[n]{z_0}=\sqrt[n]{r_0}\left(\cos\frac{\phi_0+2\pi n}{n}+i\sin\frac{\phi_0+2\pi n}{n}\right)=\sqrt[n]{r_0}\left(\cos\frac{\phi_0}{n}+i\sin\frac{\phi_0}{n}\right),$$

e si ritorna quindi al valore iniziale della radice.

Superfici di Riemann delle funzioni a più valori.

Il carattere multivoco di una funzione può essere rappresentato mediante una costruzione geometrica intuitiva.

Consideriamo ancora Ln z e tagliamo il piano z lungo il semiasse reale positivo. Impedendo a z di attraversare il taglio, non possiamo passare con continuità da un valore del logaritmo a un altro nello stesso punto. Prolungando il valore scelto in z_0 e ritornandovi senza attraversare il taglio, ritroviamo sempre lo stesso valore.

La funzione univoca ottenuta sul piano tagliato è un ramo di Ln z. Tutti i valori del logaritmo si distribuiscono in un insieme infinito di rami univoci:

$$\ln r+i\phi,\quad 2\pi n<\phi\leqslant 2\pi(n+1).$$

È facile verificare che il valore limite dell'n-esimo ramo sul bordo inferiore del taglio coincide con quello del ramo $(n+1)$-esimo sul bordo superiore.

Per distinguere i rami, immaginiamo infinite copie del piano z, ciascuna tagliata lungo il semiasse reale positivo, e associamo al foglio n i valori del ramo n-esimo. Punti di fogli diversi con le stesse coordinate rappresentano lo stesso numero $x+iy$; il foglio sul quale si trovano specifica però quale ramo del logaritmo stiamo considerando.

Per rappresentare la coincidenza dei valori ai bordi, incolliamo il bordo inferiore del taglio del foglio n al bordo superiore del foglio $n+1$. Otteniamo così una superficie a infiniti fogli, simile a una scala a chiocciola (figura 23). Il punto $z=0$, escluso dalla superficie del logaritmo, occupa idealmente il posto dell'asse centrale della scala.

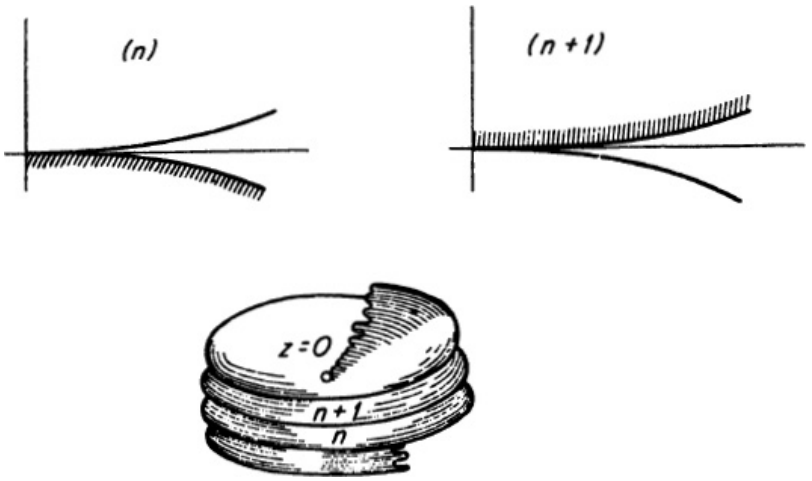


Figura 23.

Quando un punto compie un giro e passa da un foglio all'altro, il numero complesso torna al valore iniziale, ma Ln *z* passa a un altro ramo.

La superficie così costruita si chiama superficie di Riemann del logaritmo. Riemann introdusse l'idea di rappresentare mediante superfici il carattere delle funzioni analitiche a più valori e ne mostrò la fecondità.

Consideriamo anche la superficie di Riemann di $w = \sqrt{z}$. Questa funzione ha due valori e un punto di diramazione nell'origine.

Immaginiamo due copie sovrapposte del piano *z*, tagliate lungo il semiasse reale positivo. Se *z* parte da z_0, compie un giro intorno all'origine e ritorna al punto iniziale, $\sqrt{z}$ passa da un ramo all'altro; il punto sulla superficie deve quindi passare al secondo foglio. Per ottenere ciò, incolliamo il bordo inferiore del primo foglio a quello superiore del secondo. Dopo un secondo giro, $\sqrt{z}$ deve riacquistare il valore iniziale e il punto deve ritornare al primo foglio: incolliamo quindi il bordo inferiore del secondo al bordo superiore del primo. Otteniamo una superficie a due fogli. La sua rappresentazione nello spazio può mostrare un'intersezione lungo il semiasse positivo, ma questa non identifica ulteriormente i punti della superficie astratta. La figura 24 ne illustra un intorno del punto sopra $z = 0$.

Analogamente si costruiscono superfici a più fogli per rappresentare il carattere delle funzioni multivoche; i fogli sono collegati secondo il comportamento del prolungamento intorno ai punti di diramazione. Le proprietà analitiche delle funzioni risultano strettamente legate alla geometria di queste superfici. Esse non sono soltanto un ausilio illustrativo, ma uno strumento fondamentale per studiare le funzioni e sviluppare metodi d'indagine. Hanno costituito un ponte fra analisi e geometria nel campo complesso, collegando alla geometria proprietà analitiche profonde e contribuendo allo sviluppo della topologia, che studia le proprietà delle figure invarianti sotto deformazioni continue appropriate.

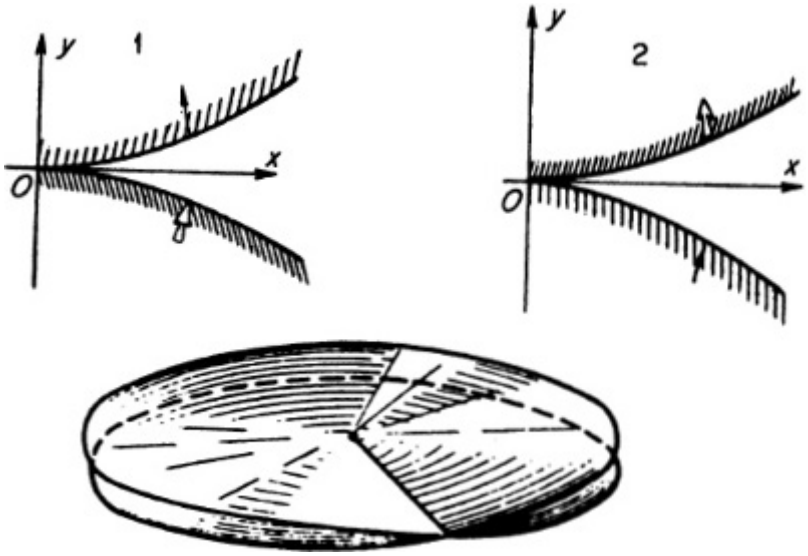


Figura 24.

Uno degli esempi più chiari è la teoria delle funzioni algebriche, definite come soluzioni di un'equazione

$$f(z, w) = 0$$

il cui primo membro è un polinomio in z e w. La superficie di Riemann connessa associata, dopo il completamento e la regolarizzazione necessari, è topologicamente una sfera oppure una sfera con manici (figura 25). Il numero dei manici è il genere della superficie e della funzione algebrica. Il genere determina molte delle proprietà più importanti della funzione.

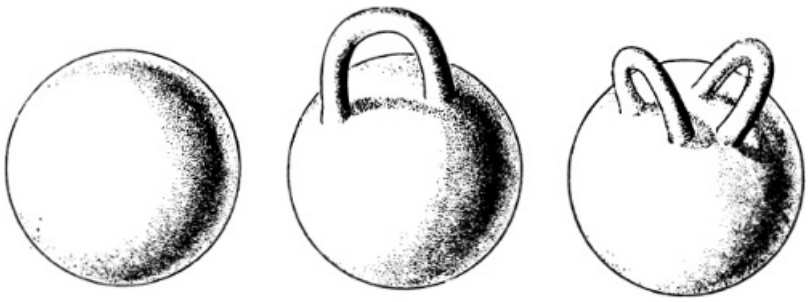


Figura 25.

§6. Conclusione

La teoria delle funzioni analitiche nacque in relazione alla soluzione delle equazioni algebriche, ma nel suo sviluppo entrò continuamente in contatto con nuovi rami della matematica. Essa chiarì le proprietà delle classi fondamentali di funzioni dell'analisi, della meccanica e della fisica matematica. Molti fatti centrali dell'analisi poterono essere compresi pienamente soltanto passando al campo complesso. Le funzioni complesse ricevettero un'interpretazione fisica immediata nei campi vettoriali dell'idrodinamica e dell'elettrodinamica, offrendo strumenti notevoli per risolverne i problemi. Si scoprirono legami con la conduzione del calore, l'elasticità e altri campi.

Le questioni generali della teoria delle equazioni differenziali e i metodi speciali per risolverle si sono sempre basati in larga misura sulle funzioni complesse. Le funzioni analitiche entrarono naturalmente nella teoria delle equazioni integrali e degli operatori lineari; si scoprirono legami profondi con la geometria. Queste connessioni, in continuo ampliamento, testimoniano la vitalità della teoria e l'arricchimento costante dei suoi problemi.

In questa rassegna non abbiamo potuto dare un quadro completo delle numerose ramificazioni della teoria. Abbiamo cercato soltanto di suggerire la varietà dei suoi problemi, indicando i fatti elementari fondamentali di alcune direzioni di ricerca. Sono rimasti esclusi aspetti importanti: i legami con le equazioni differenziali e le funzioni speciali, con le funzioni ellittiche e automorfe, con le serie trigonometriche e con molti altri rami della matematica. In altri casi abbiamo dovuto limitarci a brevi cenni. Speriamo tuttavia di aver dato al lettore un'idea generale del carattere e dell'importanza della teoria delle funzioni di variabile complessa.

Letture consigliate

R. V. Churchill, Introduction to complex variables and applications, McGraw-Hill, New York, 1948.

P. Franklin, Functions of complex variables, Prentice-Hall, Englewood Cliffs, New Jersey, 1958.

G. N. Watson, Complex integration and Cauchy's theorem, Hafner, New York, 1960.

Note al capitolo IX

[1] I numeri complessi sono noti al lettore dagli studi secondari. Si veda anche il capitolo IV, §3.

[2] Cfr. capitolo IV, §3.

[3] Cfr. capitolo III, §11.

[4] Cfr. capitolo IV, §3.

[5] Nell'uso terminologico dell'originale, il caso degenere $\Delta = 0$ non viene chiamato trasformazione affine. Per le trasformazioni affini si veda anche il capitolo III, §11.

[6] Cfr. capitolo X sulla teoria dei numeri.

Nota editoriale

La traduzione conserva le sei sezioni, le venticinque figure, le formule e le sei note del capitolo. La figura 16 comprende entrambi i disegni dell'originale. Sono corretti refusi in indici, simboli, rinvii e nella didascalia della figura 5. Nella somma geometrica il numeratore è $1 - (-1)^n z^{2n}$; nella (23) compare $\partial v/\partial y$, nella velocità del flusso lineare compare il coniugato di A, e nel limite del rapporto incrementale compare $f'(z)$. La prima formula di Taylor è riprodotta nella forma corretta presente più avanti nel capitolo, con la potenza $(z - a)^n$.

Per le serie di potenze si distingue il disco aperto di convergenza dal comportamento sulla circonferenza. Il teorema fondamentale dell'algebra conta le radici con molteplicità. L'argomento di un numero complesso è determinato modulo 2π; per un ramo analitico del logaritmo il taglio è escluso e gli estremi dell'intervallo degli argomenti sono entrambi esclusi. Le formule con l'arcotangente richiedono la corretta scelta del quadrante.

Le applicazioni ai fluidi riguardano il modello piano, stazionario, incomprimibile e irrotazionale, con le ulteriori ipotesi ideali usate nei singoli esempi. Il potenziale delle velocità può essere soltanto locale o multivoco. La circolazione ha segno dipendente dall'orientazione; nella (33) Γ positivo corrisponde al verso orario. La portanza è una forza per unità di apertura. I punti di ristagno distinti della figura 11 corrispondono al regime illustrato, non a ogni valore della circolazione. La corrispondenza con un punto di ristagno sul cerchio non implica automaticamente velocità nulla al bordo affilato del profilo. La stima storica sulla comprimibilità dell'aria è riportata come tale, senza attribuirle validità generale.

Nel teorema di rappresentazione conforme si esclude l'intero piano. I domini esterni sono semplicemente connessi soltanto includendo il punto all'infinito; nel teorema di Cauchy si usa invece la semplice connessione nel piano e l'analiticità su tutti i punti del dominio. Le descrizioni mediante ellissi sono locali e richiedono coefficienti positivi e assenza di degenerazione; per una trasformazione quasiconforme globale occorre inoltre un controllo uniforme della distorsione.

La forma del teorema di Picard con valori assunti infinite volte riguarda le funzioni intere trascendenti. Uno zero del denominatore può essere cancellato da uno del numeratore e non genera necessariamente un polo. Le parti principali determinano una funzione meromorfa soltanto a meno di una funzione intera; una loro somma infinita richiede condizioni di convergenza o termini correttivi. La serie della cotangente è raggruppata simmetricamente. Le affermazioni sul legame fra crescita e distribuzione dei valori hanno carattere introduttivo e richiedono enunciati più precisi nella teoria generale.

Il principio di identità si riferisce a un dominio connesso. Le superfici di Riemann sono superfici astratte: le intersezioni visibili nei modelli tridimensionali non costituiscono ulteriori identificazioni. Per le funzioni algebriche, la classificazione mediante sfere con manici riguarda la superficie connessa regolarizzata e completata. I riferimenti storici e allo stato delle ricerche restano quelli dell'epoca dell'opera.

CAPITOLO X

I numeri primi

K. K. Mardžanišvili e A. B. Postnikov

§1. Lo studio della teoria dei numeri

I numeri interi.

Come il lettore sa dall'introduzione al capitolo I, l'umanità ebbe a che fare con i numeri interi fin dai tempi più remoti; ma occorsero molti secoli perché si formasse il concetto della successione infinita dei numeri naturali

$$1, 2, 3, 4, 5, \cdots. \qquad (1)$$

Oggi, nelle più diverse questioni dell'attività pratica, incontriamo continuamente problemi che coinvolgono numeri interi. Questi esprimono molte relazioni quantitative della natura; in tutte le questioni riguardanti oggetti discreti costituiscono lo strumento matematico necessario.

Inoltre, i numeri interi hanno una funzione importante nello studio del continuo. Così, per esempio, nell'analisi matematica si considera lo sviluppo di una funzione analitica in una serie di potenze intere di x

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots.$$

Tutti i calcoli si eseguono essenzialmente con numeri interi, come appare subito anche da un esame superficiale delle macchine calcolatrici automatiche, delle calcolatrici da tavolo o delle tavole matematiche, per esempio quelle dei logaritmi. Compiute queste operazioni sui numeri interi, si collocano le virgole decimali in posizioni ben determinate, formando così frazioni decimali; queste, come tutte le frazioni razionali, rappresentano quozienti di due numeri interi. Quando nell'attività pratica utilizziamo un numero reale qualsiasi, per esempio π, lo sostituiamo in effetti con una frazione razionale: poniamo, per esempio, $\pi \approx 22/7$ oppure $\pi \approx 3{,}14$.

Mentre la determinazione delle regole per operare sui numeri è compito dell'aritmetica, le proprietà più profonde della successione dei numeri naturali (1), estesa a comprendere lo zero e gli interi negativi, sono studiate dalla teoria dei numeri. Essa è la scienza del sistema degli interi e, in senso più ampio, anche dei sistemi numerici costruiti in modo determinato a partire dagli interi; si veda, in particolare, il §5 di questo capitolo. È inteso che la teoria dei numeri considera gli interi nelle loro reciproche dipendenze: ne studia proprietà definite da determinate relazioni fra essi.

Una delle questioni fondamentali della teoria dei numeri riguarda la divisibilità di un numero per un altro: se il risultato della divisione dell'intero a per l'intero b, diverso da zero, è un intero, cioè se

$$a = b \cdot c$$

dove a, b, c sono interi, diciamo che a è divisibile per b, oppure che b divide a. Se il quoziente dell'intero a per l'intero b è una frazione non intera, diciamo che a non è divisibile per b. Le questioni di divisibilità si incontrano continuamente nella pratica e hanno una funzione importante anche in alcuni problemi dell'analisi matematica. Per esempio, se nello sviluppo di una funzione in potenze intere di x

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + \cdots \qquad (2)$$

tutti i coefficienti di indice dispari, cioè di indice non divisibile per 2, sono nulli, ossia se

$$f(x) = a_0 + a_2x^2 + \cdots + a_{2k}x^{2k} + \cdots,$$

la funzione soddisfa la condizione

$$f(-x) = f(x);$$

Una funzione di questo tipo si chiama pari, e il suo grafico è simmetrico rispetto all'asse delle ordinate. Se invece nello sviluppo (2) sono nulli tutti i coefficienti di indice pari, cioè di indice divisibile per 2, in altre parole se

$$f(x) = a_1x + a_3x^3 + \cdots + a_{2k+1}x^{2k+1} + \cdots,$$

allora

$$f(-x) = -f(x);$$

In questo caso la funzione si chiama dispari, e il suo grafico è simmetrico rispetto all'origine.

Così, per esempio,

$$\sin x = x - x^3/3! + x^5/5! - \ldots \text{ (funzione dispari);}$$
$$\cos x = 1 - x^2/2! + x^4/4! - \ldots \text{ (funzione pari).}$$

La questione geometrica della possibilità di costruire con riga e compasso un poligono regolare di n lati risulta dipendere dalla natura aritmetica del numero n.[1]

Si chiama primo un numero intero maggiore di uno che possiede soltanto due divisori interi positivi: uno e il numero stesso. Uno non è considerato primo, perché non possiede due divisori positivi distinti.

I numeri primi sono dunque

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \cdots. \qquad (3)$$

I numeri primi hanno un ruolo fondamentale nella teoria dei numeri per il seguente teorema: ogni intero $n > 1$ si può rappresentare come prodotto di numeri primi, con eventuali ripetizioni dei fattori, cioè nella forma

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}, \qquad (4)$$

dove $p_1 < p_2 < \ldots < p_k$ sono primi e $a_1, a_2, \ldots, a_k$ sono interi non minori di uno. Inoltre, la rappresentazione di n nella forma (4) è unica.

Le proprietà dei numeri collegate alla loro rappresentazione come somma di termini si chiamano additive; quelle relative alla rappresentazione come prodotto si chiamano moltiplicative. Il legame fra proprietà additive e moltiplicative è straordinariamente complesso e ha dato origine a una serie di problemi fondamentali della teoria dei numeri.

L'esistenza di questi difficili problemi, insieme al fatto che il numero intero è il più semplice e chiaro dei concetti matematici ed è strettamente legato alla realtà oggettiva, ha condotto alla creazione di idee profonde e di metodi potenti per la teoria dei numeri; molti di essi hanno acquistato importanza anche in altri rami della matematica. Per esempio, l'idea della successione infinita dei numeri naturali, che riflette l'infinità del mondo materiale nello spazio e nel tempo, ha esercitato un'enorme influenza sull'intero sviluppo della matematica. Ha grande significato anche il fatto che i termini di questa successione siano ordinati. Lo studio delle operazioni sugli interi ha condotto al concetto di operazione algebrica, che svolge un ruolo fondamentale in diversi rami della matematica.

Un'importanza immensa ha avuto in matematica il concetto di algoritmo, particolarmente adatto alle questioni aritmetiche: un procedimento per risolvere problemi mediante la ripetizione di operazioni rigorosamente definite. Il ruolo degli algoritmi è fondamentale, in particolare, nell'uso delle macchine matematiche. La natura essenziale del metodo algoritmico è illustrata chiaramente dall'algoritmo di Euclide per trovare il massimo comune divisore di due numeri naturali a e b.

Supponiamo $a > b$. Dividiamo a per b e troviamo il quoziente q_1 e, se b non divide a, il resto r_2:

$$a = bq_1 + r_2, \quad 0 < r_2 < b. \qquad (5_1)$$

Poi, se $r_2 \neq 0$, dividiamo b per r_2:

$$b = r_2 q_2 + r_3, \quad 0 < r_3 < r_2. \qquad (5_2)$$

Dividiamo quindi r_2 per r_3 e continuiamo fino a ottenere un resto nullo: ciò deve necessariamente avvenire, perché i resti non nulli $r_2, r_3, \ldots$ formano una successione decrescente di interi positivi. Sia

$$r_{n-2} = r_{n-1} q_{n-1} + r_n, \qquad (5_{n-1})$$

$$r_{n-1} = r_n q_n, \qquad (5_n)$$

Allora si vede subito che r_n è il massimo comune divisore di a e b. Infatti, se due interi l e m hanno un divisore comune d, per qualsiasi coppia di interi h e k anche $hl + km$ è divisibile per d. Indichiamo con δ il massimo comune divisore di a e b. Dalla (5_1) si vede che δ divide r_2; dalla (5_2) segue che divide r_3, ...; dalla (5_{n-1}) che divide r_n. Ma r_n è a sua volta un divisore comune di a e b: dalla (5_n) risulta che divide r_{n-1}, dalla (5_{n-1}) che divide r_{n-2}, e così via, risalendo fino ad a e b. Dunque δ coincide con r_n, e il problema è risolto. Abbiamo un procedimento ben definito, dello stesso tipo per ogni a e b, che conduce automaticamente al risultato voluto: un esempio caratteristico di algoritmo.

La teoria dei numeri ha influenzato lo sviluppo di molte discipline matematiche: l'analisi, la geometria, l'algebra classica e contemporanea, la teoria della sommabilità delle serie, il calcolo delle probabilità e così via.

I metodi della teoria dei numeri.

Dal punto di vista dei metodi, la teoria dei numeri si divide in quattro parti: elementare, analitica, algebrica e geometrica.

La teoria elementare dei numeri studia le proprietà degli interi senza ricorrere ad altre discipline matematiche. Così, partendo dall'identità di Eulero

$$\begin{aligned}(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) &= (x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4)^2 \\ &+ (x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3)^2 + (x_1y_3 - x_3y_1 + x_4y_2 - x_2y_4)^2 \\ &+ (x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2)^2, \qquad (6)\end{aligned}$$

si può dimostrare in modo assai semplice che ogni intero $N > 0$ si può esprimere come somma dei quadrati di quattro interi, cioè nella forma

$$N = x^2 + y^2 + z^2 + u^2,$$

dove x, y, z e u sono interi.[2]

La teoria analitica dei numeri applica l'analisi matematica ai problemi della teoria dei numeri. Le sue basi furono poste da Eulero; essa fu sviluppata da P. L. Čebyšev, Dirichlet, Riemann, Ramanujan, Hardy, Littlewood e altri matematici, mentre i suoi metodi più potenti sono dovuti a Vinogradov. Questa parte della teoria dei numeri è strettamente connessa con la teoria delle funzioni di variabile complessa, ricchissima di applicazioni pratiche, e anche con la teoria delle serie, il calcolo delle probabilità e altri rami della matematica.

Il concetto fondamentale della teoria algebrica dei numeri è quello di numero algebrico, cioè di radice di un'equazione

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_{n-1}x + a_n = 0,$$

dove a_0, a_1, a_2, ..., a_n sono interi e $a_0 \neq 0$.[3]

I maggiori contributi a questo ramo della teoria dei numeri sono dovuti a Lagrange, Gauss, Kummer, E. I. Zolotarev, Dedekind, A. O. Gel'fond e altri.

Gli oggetti fondamentali della teoria geometrica dei numeri sono i «reticoli spaziali»: sistemi costituiti da tutti i punti «interi», le cui coordinate, in un dato sistema di coordinate rettilinee, ortogonali oppure oblique, sono numeri interi. I reticoli hanno grande importanza nella geometria e nella cristallografia e sono intimamente connessi con questioni rilevanti della teoria dei numeri, in particolare con la teoria aritmetica delle forme quadratiche, ossia delle forme quadratiche con coefficienti interi e variabili intere. I lavori fondamentali della teoria geometrica dei numeri sono dovuti a H. Minkowski e G. F. Voronoj.

Occorre osservare che i metodi della teoria analitica dei numeri trovano applicazioni importanti anche nei due rami algebrico e geometrico. Merita particolare attenzione il problema di contare i punti interi contenuti in un dato dominio, importante in alcuni settori della fisica. Diversi modi di affrontarlo furono indicati da G. F. Voronoj, mentre I. M. Vinogradov sviluppò metodi per risolverlo.

La ragione profonda della potenza dei metodi analitici nella teoria dei numeri è che essi arricchiscono lo studio delle relazioni fra interi discreti, facendo intervenire nuove relazioni fra grandezze continue.

Dobbiamo sottolineare che in questo capitolo prenderemo in esame soltanto alcune questioni scelte della teoria dei numeri.

§2. Le ricerche sui numeri primi

I numeri primi sono infiniti.

Considerando la successione (3) dei numeri primi

$$2, 3, 5, 7, 11, 13, 17, 19, \cdots$$

è naturale domandarsi se essa sia infinita. Il fatto che ogni intero maggiore di uno si possa rappresentare nella forma (4) non risolve ancora il problema, perché gli esponenti $a_1, \ldots, a_k$ possono assumere infiniti valori. Una risposta affermativa fu data da Euclide, che dimostrò che il numero dei primi non può essere uguale ad alcun intero finito k.

Siano $p_1, p_2, \ldots, p_k$ numeri primi. Allora il numero

$$m = p_1 p_2 \ldots p_k + 1,$$

essendo un intero maggiore di uno, o è primo oppure ha un fattore primo. Ma m non è divisibile per nessuno dei primi $p_1, \ldots, p_k$: se lo fosse, sarebbe divisibile per quel primo anche la differenza $m - p_1 p_2 \ldots p_k$, il che è impossibile, perché tale differenza è uno. Dunque o m è primo, oppure è divisibile per un primo p_{k+1} distinto da $p_1, \ldots, p_k$. L'insieme dei primi non può quindi essere finito.

Il crivello di Eratostene.

Nel III secolo a.C. il matematico greco Eratostene descrisse il seguente metodo del «crivello» per trovare tutti i primi non maggiori di un dato numero naturale *N*. Scriviamo tutti gli interi da 1 a *N*:

$$1, 2, 3, 4, \cdots, N,$$

Poi, procedendo da sinistra, cancelliamo prima il numero 1, quindi tutti i multipli di 2 tranne il 2 stesso, tutti i multipli di 3 tranne il 3, tutti i multipli di 5 tranne il 5 - i multipli di quattro sono già stati cancellati - e così via. I numeri rimasti sono primi. È notevole che basti proseguire le cancellazioni fino ad aver trattato tutti i primi non maggiori di $\sqrt{N}$, poiché ogni numero composto, cioè non primo, non maggiore di *N* possiede necessariamente un divisore primo non maggiore di $\sqrt{N}$.

L'esame della successione dei primi all'interno di quella degli interi positivi induce a pensare che la loro legge di distribuzione sia molto complessa. Incontriamo, per esempio, primi come 8 004 119 e 8 004 121, detti primi gemelli, che differiscono di due, ma anche primi lontani fra loro, come 86 629 e 86 677, fra i quali non vi è alcun altro primo. Le tavole mostrano tuttavia che, «in media», i primi si incontrano sempre più raramente mentre si percorre la successione degli interi.

L'identità di Eulero e la sua dimostrazione dell'infinità dei primi.

Il grande matematico del XVIII secolo L. Eulero, membro dell'Accademia russa delle scienze, introdusse la seguente funzione dell'argomento $s > 1$, che oggi si indica con $\zeta(s)$:

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} + \cdots . \qquad (7)$$

Come sappiamo dal capitolo II, questa serie converge per $s > 1$ e diverge per $s \leqslant 1$. Eulero ricavò una notevole identità, che ha un ruolo importantissimo nella teoria dei numeri primi:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p} \frac{1}{1 - \frac{1}{p^s}}, \qquad (8)$$

dove il simbolo Π_p indica che si devono moltiplicare le espressioni $1/[1 - (1/p^s)]$ per tutti i primi *p*. Per comprendere la dimostrazione di questa identità, osserviamo che $1/(1 - q) = 1 + q + q^2 + \ldots$ per $|q| < 1$, cosicché

$$\frac{1}{1 - \frac{1}{p^s}} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots .$$

Moltiplicando queste serie per i diversi primi *p* e ricordando che ogni *n* si rappresenta in modo unico come prodotto di primi, troviamo

$$\prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \cdots\right) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \cdots + \frac{1}{n^s} + \cdots .$$

Naturalmente, per una dimostrazione rigorosa occorre giustificare il passaggio al limite, ma ciò non presenta particolari difficoltà.

Dall'identità (8) si può ricavare, come corollario, che la serie $\Sigma_p\ 1/p$, formata dai reciproci di tutti i primi, diverge. Ciò fornisce una nuova dimostrazione del fatto, già noto, che i primi non possono essere in numero finito. Si ricava inoltre che il rapporto fra il numero dei primi non maggiori di x e x stesso tende a zero quando x cresce indefinitamente.

Le ricerche di Čebyšev sulla distribuzione dei primi.

Indichiamo con $\pi(x)$, secondo l'uso ormai corrente, il numero dei primi non maggiori di x. Per esempio, $\pi(10) = 4$, perché 2, 3, 5 e 7 sono tutti e soli i primi non maggiori di 10; $\pi(\pi) = 2$, perché 2 e 3 sono tutti e soli i primi non maggiori di π. Come si è osservato,

$$\lim_{x\to\infty} \frac{\pi(x)}{x} = 0.$$

Ma come diminuisce precisamente il rapporto $\pi(x)/x$? In altre parole, qual è la legge di crescita di $\pi(x)$? Possiamo cercare una funzione abbastanza semplice e ben nota che differisca poco da $\pi(x)$? Il celebre matematico francese Legendre, studiando le tavole dei numeri primi, affermò che una tale funzione sarebbe stata

$$\frac{x}{\ln x - A}, \qquad (9)$$

dove $A = 1{,}08\ldots$, ma non diede una dimostrazione. Gauss, che si occupò anch'egli della distribuzione dei primi, congetturò che $\pi(x)$ differisse relativamente poco da $\int_2^x dt/\ln t$. Osserviamo che vale la relazione

$$\lim_{x\to\infty} \frac{\int_2^x \frac{dt}{\ln t}}{\frac{x}{\ln x}} = 1, \qquad (10)$$

che si stabilisce integrando per parti e stimando il nuovo integrale.

Il primo matematico, dopo Euclide, a compiere un progresso reale nella difficilissima questione della distribuzione dei primi fu P. L. Čebyšev. Nel 1848, basandosi sullo studio della funzione $\zeta(s)$ di Eulero per s reale, Čebyšev dimostrò che, per ogni intero positivo n, arbitrariamente grande, e per ogni α positivo, arbitrariamente piccolo, esistono valori di x arbitrariamente grandi per i quali

$$\pi(x) > \int_2^x \frac{dt}{\ln t} - \frac{\alpha x}{\ln^n x},$$

ed esistono anche valori di x arbitrariamente grandi per i quali

$$\pi(x) < \int_2^x \frac{dt}{\ln t} + \frac{\alpha x}{\ln^n x},$$

Ciò concorda bene con l'ipotesi di Gauss. In particolare, ponendo $n = 1$ e applicando la (10), Čebyšev stabilì che

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\ln x}} = 1, \qquad (11)$$

purché il limite nella (11) esista.

Čebyšev confutò inoltre l'ipotesi di Legendre sul valore della costante A nell'espressione (9), intesa come migliore approssimazione di $\pi(x)$: dimostrò che tale valore può essere soltanto $A = 1$.

Le ricerche sulla teoria dei gruppi condussero il noto matematico francese Bertrand alla seguente congettura, che egli verificò empiricamente mediante le tavole fino a valori di n piuttosto grandi: se $n > 3$, fra n e $2n - 2$ esiste almeno un primo. Tutti i tentativi di Bertrand e di altri matematici per dimostrarla rimasero infruttuosi fino al 1850, quando Čebyšev pubblicò il suo secondo articolo sui numeri primi. In esso dimostrò questa congettura, detta «postulato di Bertrand», e inoltre provò che, per x sufficientemente grande,

$$A_1 < \frac{\pi(x)}{\frac{x}{\ln x}} < A_2, \qquad (12)$$

dove

$$0{,}92 < A_1 < 1 \text{ e } 1 < A_2 < 1{,}1.$$

Nel §3 presenteremo una versione semplificata del metodo di Čebyšev, che conduce però a risultati sensibilmente meno precisi di quelli ottenuti da lui.

I lavori di Čebyšev ebbero una grande influenza su molti matematici, in particolare su Sylvester e Poincaré. Per oltre quarant'anni numerosi studiosi si occuparono di migliorare la disuguaglianza (12), aumentando la costante a sinistra e diminuendo quella a destra; non riuscirono tuttavia a stabilire l'esistenza del limite

$$\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\ln x}}$$

Come si è già osservato, sappiamo dal lavoro di Čebyšev che, se questo limite esiste, è uguale a uno.

Soltanto nel 1896 Hadamard, ricorrendo alla teoria delle funzioni di variabile complessa, dimostrò che la funzione $\Theta(x)$, introdotta da Čebyšev e definita dall'equazione

$$\Theta(x) = \sum_{p \leqslant x} \ln p,$$

soddisfa la condizione

$$\lim_{x \to \infty} \frac{\Theta(x)}{x} = 1, \qquad (13)$$

da cui è relativamente facile ottenere la relazione (11) senza ulteriori ipotesi: è la cosiddetta legge asintotica della distribuzione dei numeri primi.

Hadamard ottenne il risultato (13) sulla base delle ricerche del celebre matematico tedesco del XIX secolo Riemann, che studiò la funzione $\zeta(s)$ di Eulero (7) per valori complessi della variabile $s = \sigma + it$. Čebyšev aveva invece considerato questa funzione soltanto per valori reali dell'argomento.[4]

Riemann dimostrò che la funzione $\zeta(s)$, definita nel semipiano $\sigma > 1$ dalla serie (7)

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

ha la proprietà che

$$\zeta(s) - \frac{1}{s-1}$$

è una funzione intera trascendente. Per $\sigma \leqslant 1$ la serie (7) non converge più, ma i valori di $\zeta(s)$ nel semipiano $\sigma \leqslant 1$ sono definiti mediante il prolungamento analitico, con l'eccezione del polo in $s = 1$; si veda il capitolo IX. Riemann formulò la congettura, detta «ipotesi di Riemann», che tutte le radici di $\zeta(s)$ nella striscia $0 \leqslant \sigma \leqslant 1$ abbiano parte reale uguale a $\frac{1}{2}$, cioè giacciano sulla retta $\sigma = \frac{1}{2}$. La questione della validità di questa ipotesi rimane aperta all'epoca dell'opera.

Un passo importante nella dimostrazione della (13) fu stabilire che sulla retta $\sigma = 1$ non vi sono radici di $\zeta(s)$.

Lo studio del comportamento di $\zeta(s)$ condusse allo sviluppo di un'elegante teoria delle funzioni intere e meromorfe, con importanti applicazioni pratiche.

Il lavoro di Vinogradov e dei suoi allievi sui numeri primi.

Dall'equazione (13), che in virtù della (10) si può scrivere nella forma

$$\lim_{x\to\infty} \frac{\pi(x)}{\int_2^x \frac{dt}{\ln t}} = 1, \qquad (14)$$

nacque la questione della precisione con cui la funzione $\int_2^x dt/\ln t$ rappresenta $\pi(x)$. I migliori risultati in questa direzione furono ottenuti da N. G. Čudakov sulla base del metodo delle somme trigonometriche di Vinogradov, che sarà descritto nel §4. Questo metodo consentì a Čudakov anche di restringere notevolmente gli intervalli entro i quali si può trovare almeno un primo. Si era stabilito in precedenza che, considerando la successione

$$1^{250}, 2^{250}, 3^{250}, \cdots, n^{250}, (n+1)^{250}, \cdots, \qquad (15)$$

a partire da un certo $n = n_0$, fra due termini consecutivi qualsiasi, cioè fra n^{250} e $(n+1)^{250}$, deve esistere almeno un primo.

Osserviamo che, come segue dalla formula del binomio,

$$(n+1)^{250} - n^{250} > 250n^{249},$$

questa differenza è molto grande. N. G. Čudakov riuscì a sostituire alla successione (15) la successione

$$1^4, 2^4, 3^4, \cdots, n^4, (n+1)^4, \cdots, \qquad (16)$$

i cui termini sono molto più vicini fra loro, ma che contiene anch'essa almeno un primo fra ogni coppia di termini consecutivi, cioè fra n^4 e $(n+1)^4$, a partire da un certo $n = n_0$. In seguito il risultato fu migliorato sostituendo le quarte potenze con i cubi.

Se k e l sono primi fra loro, cioè non hanno divisori comuni maggiori di uno, una progressione aritmetica di termine generale $kt + l$, con k positivo, contiene infiniti numeri primi. Questo fatto, che generalizza il risultato di Euclide, fu stabilito nel XIX secolo da Dirichlet. Ma possiamo trovare un limite che certamente non sia superato dal più piccolo primo della progressione? Il matematico di Leningrado Ju. V. Linnik dimostrò l'esistenza di una costante assoluta C tale che, per $k > 1$ e per una classe residua l prima con k, esiste almeno un primo minore di k^C nella corrispondente progressione. Linnik fornì così una soluzione sostanzialmente completa del problema, posto molti anni prima, del più piccolo primo di una progressione aritmetica; le ricerche successive possono diminuire il valore della costante C. Linnik condusse inoltre ricerche importantissime sugli zeri della funzione $\zeta(s)$ e di funzioni più generali.

Come si è già ricordato, i migliori risultati sulla distribuzione dei primi furono ottenuti mediante il metodo di Vinogradov per stimare somme trigonometriche.

Una somma trigonometrica è una somma della forma

$$\sum_{A<x<B} e^{2\pi i f(x)},$$

dove $f(x)$ è una funzione reale di x, e x assume tutti i valori interi compresi fra A e B, oppure soltanto quelli di un sottoinsieme specificato, per esempio i primi fra A e B. Poiché il modulo di $e^{2\pi iz}$, per z reale, è uguale a uno, e il modulo di una somma non supera la somma dei moduli dei suoi termini, abbiamo

$$\left| \sum_{x=1}^{p} e^{2\pi i f(x)} \right| \leqslant P. \qquad (17)$$

Nella (17) si è scritto il caso di una somma su P interi consecutivi; in generale il limite è il numero dei termini. Questa stima «banale» può essere migliorata notevolmente in numerosi casi; i passi decisivi in questa direzione furono compiuti da Vinogradov. Per fissare le idee, sia $f(x)$ un polinomio

$$f(x) = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \cdots + \alpha_1 x + \alpha_0 .$$

Se tutti gli α sono interi, allora $e^{2\pi i f(x)} = 1$ per x intero, e in questo caso la stima (17) evidentemente non si può migliorare. Ma se $\alpha_1, \ldots, \alpha_n$ non sono tutti interi, come dimostrò Vinogradov, la stima può in numerosi casi essere affinata approssimando questi coefficienti con frazioni razionali aventi denominatori non superiori a un certo limite. Si può dimostrare che ogni α fra 0 e 1 è rappresentabile nella forma $\alpha = a/q + z$, dove a e q sono interi primi fra loro, $1 \leq q \leq \tau$, $|z| \leq 1/(q\tau)$ e τ è un intero prefissato maggiore di 1.

La creazione del metodo delle somme trigonometriche permise a Vinogradov di risolvere una serie di difficilissimi problemi della teoria dei numeri. In particolare, nel 1937 egli risolse, per gli interi dispari sufficientemente grandi, un celebre problema posto da Goldbach: dimostrò che ogni tale intero N si può rappresentare come somma di tre primi,

$$N = p_1 + p_2 + p_3 . \qquad (18)$$

Il problema era sorto nel 1742 nella corrispondenza fra Eulero e un altro membro dell'Accademia russa delle scienze, C. Goldbach, ed era rimasto irrisolto per quasi due secoli, nonostante gli sforzi di numerosi matematici eminenti.

Come abbiamo visto, la (4) mostra il ruolo fondamentale dei primi nella rappresentazione moltiplicativa di un numero; la (18) ne fornisce invece una rappresentazione additiva per gli interi dispari sufficientemente grandi. Dalla (18) è facile mostrare che un numero pari sufficientemente grande si può rappresentare come somma di non più di quattro primi.[5] Il teorema di Vinogradov-Goldbach stabilì così un legame profondo fra proprietà additive e moltiplicative dei numeri.

Il significato del metodo delle somme trigonometriche creato da Vinogradov non si limita alla teoria dei numeri. Esso svolge un ruolo importante, in particolare, nella teoria delle funzioni e nel calcolo delle probabilità. Un'idea del metodo si può ricavare dal §4 di questo capitolo.

Il lettore interessato a una trattazione più approfondita può consultare il libro di Vinogradov «Il metodo delle somme trigonometriche nella teoria dei numeri», dopo una lettura preliminare del suo volume «Fondamenti della teoria dei numeri».

§3. Il metodo di Čebyšev

La funzione ϑ di Čebyšev e le sue stime. Presentiamo ora una versione semplificata del metodo di Čebyšev per calcolare il numero dei primi compresi entro limiti assegnati. Per brevità adottiamo la seguente notazione: sia B una quantità variabile positiva che può crescere indefinitamente, e sia A un'altra quantità tale che $|A|$ cresca «non più rapidamente» di CB, dove C è una costante positiva. Più precisamente, se esiste una costante $C > 0$ tale che, da un certo punto in poi, valga sempre $|A|/B \leqslant C$, scriveremo

$$A = O(B).$$

Questa scrittura si legge abitualmente: «A è una quantità dell'ordine di B». Per esempio,

$$\sin x = O(1),$$

poiché dappertutto

$$\frac{|\sin x|}{1} \leqslant 1;$$

e, esattamente allo stesso modo,

$$5x^3 \cos 2x = O(x^3).$$

Indicheremo inoltre con $[x]$ la parte intera di x, cioè il massimo intero non maggiore di x; così, per esempio,

$$[\pi] = 3,\ [5] = 5,\ [-1.5] = -2,\ [0.999] = 0.$$

Poniamoci ora la seguente domanda. Siano p un primo e n un numero naturale; indichiamo come al solito con $n!$ il prodotto $1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$. Osserviamo incidentalmente che, all'aumentare di n, il valore di $n!$ cresce molto rapidamente. Qual è il massimo esponente a per cui p^a divide $n!$ senza resto?

Fra i numeri 1, 2, ..., n, quelli divisibili per p sono precisamente $[n/p]$; fra questi, quelli divisibili anche per p^2 sono $[n/p^2]$; fra questi ultimi, $[n/p^3]$ sono divisibili per p^3, e così via. È quindi facile mostrare che

$$a = \left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots$$

La serie si arresta, perché $[n/p^s] > 0$ soltanto per $n \geqslant p^s$. Nell'ultima somma, infatti, ogni fattore del prodotto $1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ per il quale la massima potenza di p che lo divide è p^m viene contato precisamente m volte: una come multiplo di p, una come multiplo di p^2, una come multiplo di p^3, ..., e infine una come multiplo di p^m.

Da questo risultato e dalla rappresentabilità di ogni naturale maggiore di uno nella forma (4) segue che $n!$ è il prodotto delle potenze della forma

$$p^{[\frac{n}{p}]+[\frac{n}{p^2}]+[\frac{n}{p^3}]+\cdots},$$

considerate per tutti i primi $p \leqslant n$. Pertanto $\ln(n!)$ è la somma dei logaritmi di queste potenze, che si può scrivere sinteticamente nella forma

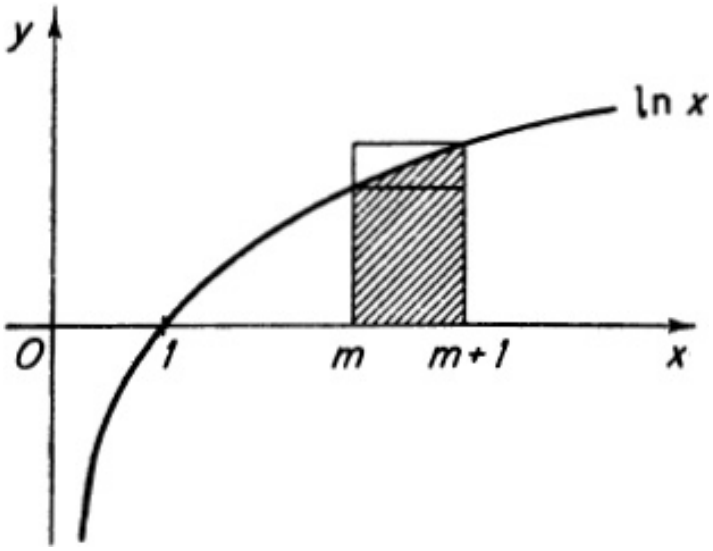


Figura 1.

$$\ln(n!) = \Sigma_{p \leq n} ([n/p] + [n/p^2] + [n/p^3] + \ldots) \ln p. \quad (19)$$

Semplifichiamo la (19). Poiché $y = \ln x$ è una funzione crescente, abbiamo

$$\ln m = \ln m \int_m^{m+1} dx < \int_m^{m+1} \ln x \, dx < \ln(m+1) \int_m^{m+1} dx = \ln(m+1)$$

come risulta chiaramente dalla figura 1. Dunque

$$\begin{aligned} \ln n! &= \ln 1 + \ln 2 + \cdots + \ln n \\ &< \int_1^2 \ln x \, dx + \int_2^3 \ln x \, dx + \cdots \\ &\quad + \int_{n-1}^n \ln x \, dx + \ln n = \int_1^n \ln x \, dx + \ln n, \end{aligned}$$

D'altra parte,

$$\ln n! > \ln 1 + \int_1^2 \ln x\,dx + \cdots + \int_{n-2}^{n-1} \ln x\,dx + \int_{n-1}^{n} \ln x\,dx = \int_1^n \ln x\,dx.$$

Usando la formula d'integrazione per parti, troviamo

$$\int_1^n \ln x\,dx = [x \cdot \ln x]_1^n - \int_1^n x \cdot \frac{1}{x} \cdot dx = n \ln n - (n-1).$$

Pertanto

$$n \ln n - n + 1 < \ln n! < n \ln n - n + 1 + \ln n,$$

e ne segue che

$$\ln n! = n \ln n + O(n). \tag{20}$$

Osserviamo che $\ln n = O(n)$; inoltre, per $n \to \infty$, la funzione $\ln n$ cresce più lentamente di qualsiasi potenza positiva di n: per ogni costante $\alpha > 0$,

$$\lim_{n\to\infty} \frac{\ln n}{n^\alpha} = 0, \tag{21}$$

poiché, in base alla regola per le forme indeterminate, si veda il capitolo II,

$$\lim_{n\to\infty} \frac{\ln n}{n^\alpha} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\alpha n^{\alpha-1}} = \frac{1}{\alpha} \lim_{n\to\infty} \frac{1}{n^\alpha} = 0.$$

Troviamo inoltre

$$\sum_{p \leqslant n} \left(\left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots \right) \ln p \leqslant \sum_{p \leqslant n} \left(\frac{n}{p^2} + \frac{n}{p^3} + \cdots \right) \ln p$$

$$= \sum_{p \leqslant n} \frac{n \ln p}{p^2 \left(1 - \frac{1}{p}\right)} < 2n \sum_{p \leqslant n} \frac{\ln p}{p^2} < 2n \sum_{m=1}^{\infty} \frac{\ln m}{m^2} = 2nC_0 = O(n), \tag{22}$$

dove C_0 è la somma della serie convergente $\sum_{m=1}^{\infty} \frac{\ln m}{m^2}$. La convergenza assoluta di questa serie si stabilisce usando la (21), per esempio con $\alpha = \frac{1}{2}$, mediante il criterio del confronto e il cosiddetto criterio integrale di convergenza; si veda il capitolo II, §14. Tenendo conto delle (20) e (22), la (19) si può porre nella forma

$$\sum_{p \leqslant n} \left[\frac{n}{p}\right] \ln p = n \ln n + O(n). \tag{23}$$

Consideriamo ora la funzione introdotta da Čebyšev

$$\Theta(n) = \sum_{p \leqslant n} \ln p \qquad (24)$$

ossia il logaritmo del prodotto di tutti i primi non maggiori di n.

La (23) si può riscrivere così:

$$\Theta\left(\frac{n}{1}\right) + \Theta\left(\frac{n}{2}\right) + \Theta\left(\frac{n}{3}\right) + \Theta\left(\frac{n}{4}\right) + \cdots = n \ln n + O(n). \qquad (25)$$

Infatti, ogni dato $\ln p$ compare in tutte le somme della forma $\Theta(n/s)$, dove $p \leqslant n/s$, cioè dove $s \leqslant n/p$,; il numero di tali somme $\Theta(n/s)$ è uguale a $[n/p]$.

La (25) vale anche quando l'argomento non è intero. Per vederlo, basta evidentemente provarla per tutti gli x tali che $n < x < n + 1$, con n intero. A questo scopo basta mostrare che sostituire n con x nel primo membro della (25) non lo modifica, mentre il primo termine del secondo membro può aumentare soltanto di una quantità $O(n)$. Il primo fatto segue perché tale sostituzione non aumenta nessun termine del primo membro: un aumento può avvenire solo quando l'argomento raggiunge il successivo intero. Naturalmente il primo membro non diminuisce. Il secondo fatto segue dalla formula dell'incremento di una funzione, si veda il capitolo II,

$$f(x) - f(a) = (x - a) f'(\xi), \quad a < \xi < x,$$

da cui abbiamo

$$x \ln x - n \ln n = (x - n) \cdot (\ln \xi + 1), \quad n < \xi < x,$$

Il secondo membro di quest'ultima equazione è minore di $\ln(n + 1) + 1 = O(n)$, poiché $0 < x - n < 1$. Sottraiamo ora dalla (25) il doppio dell'equazione ottenuta sostituendo n con $n/2$:

$$\Theta\left(\frac{n}{1}\right) + \Theta\left(\frac{n}{2}\right) + \Theta\left(\frac{n}{3}\right) + \Theta\left(\frac{n}{4}\right) + \cdots = n \ln n + O(n),$$

$$2\Theta\left(\frac{n}{2}\right) + 2\Theta\left(\frac{n}{4}\right) + \cdots = 2 \cdot \frac{n}{2} \cdot \ln \frac{n}{2} + O(n),$$

Otteniamo

$$\Theta\left(\frac{n}{1}\right) - \Theta\left(\frac{n}{2}\right) + \Theta\left(\frac{n}{3}\right) - \Theta\left(\frac{n}{4}\right) + \cdots = n \ln 2 + O(n) < C_1 n,$$

dove C_1 è una costante positiva. Ma $\Theta(n/1) - \Theta(n/2)$ non supera l'intero primo membro, perché le differenze $\Theta(n/3) - \Theta(n/4)$, $\Theta(n/5) - \Theta(n/6)$, ... non possono essere negative. Dall'ultima disuguaglianza segue dunque

$$\Theta\left(\frac{n}{1}\right) - \Theta\left(\frac{n}{2}\right) < C_1 n.$$

Sostituendo qui n con $n/2$, $n/4$, ..., otteniamo anche

$$\Theta\left(\frac{n}{2}\right) - \Theta\left(\frac{n}{4}\right) < C_1 \cdot \frac{n}{2},$$

$$\Theta\left(\frac{n}{4}\right) - \Theta\left(\frac{n}{8}\right) < C_1 \cdot \frac{n}{4},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots,$$

Poiché $\Theta(n/2^k) = 0$ per k sufficientemente grande, quando $n/2^k < 2$, sommando termine a termine si trova

$$\Theta(n) < C_1\left(n + \frac{n}{2} + \frac{n}{4} + \cdots\right) = 2C_1 n. \qquad (26)$$

Tornando alla (23), troviamo

$$0 \leqslant \sum_{p \leqslant n} \frac{n}{p} \ln p - \sum_{p \leqslant n} \left[\frac{n}{p}\right] \ln p \leqslant \sum_{p \leqslant n} \ln p = \Theta(n) \leqslant 2C_1 n = O(n),$$

e quindi la (23) dà

$$\sum_{p \leqslant n} \frac{n}{p} \ln p = n \ln n + O(n),$$

$$\sum_{p \leqslant n} \frac{\ln p}{p} = \ln n + \theta C, \qquad (27)$$

dove C è una costante positiva e θ dipende dal numero n in modo che $|\theta| \leqslant 1$..

Una stima del numero dei primi in un intervallo assegnato.

Mostriamo ora che si può scegliere una costante positiva M tale che, fra n e Mn, si trovino tanti primi p quanti se ne vogliono, purché n sia sufficientemente grande. Stabiliremo precisamente alcune semplici disuguaglianze per il numero T dei primi nell'intervallo $n < p \leqslant Mn$. Evidentemente,

$$\sum_{n < p \leqslant Mn} \frac{\ln p}{p} = \sum_{p \leqslant Mn} \frac{\ln p}{p} - \sum_{p \leqslant n} \frac{\ln p}{p}. \qquad (28)$$

Sostituendo n con Mn nella (27), otteniamo

$$\sum_{p \leqslant Mn} \frac{\ln p}{p} = \ln(Mn) + \theta' C = \ln M + \ln n + \theta' C, \qquad (29)$$

dove $|\theta'| \leqslant 1$. Quindi, tenendo conto delle (28), (29) e (27), abbiamo

$$\sum_{n<p\leqslant Mn} \frac{\ln p}{p} = \ln M + \theta' C - \theta C = \ln M + 2\theta_0 C,$$

dove $|\theta_0| \leqslant 1$, ossia

$$\ln M - 2C \leqslant \sum_{n<p\leqslant Mn} \frac{\ln p}{p} \leqslant \ln M + 2C. \qquad (30)$$

D'altra parte, poiché $y = \ln x/x$ è decrescente per $x > e$ - infatti $y' = (1 - \ln x)/x^2 < 0$ per $\ln x > 1$, cioè per $x > e$ - segue che per $n \geqslant 3$

$$T\frac{\ln Mn}{Mn} \leqslant \sum_{n<p\leqslant Mn} \frac{\ln p}{p} \leqslant T\frac{\ln n}{n},$$

Pertanto dalla (30) abbiamo

$$T\frac{\ln n}{n} > \ln M - 2C \qquad (31)$$

e

$$T\frac{\ln (Mn)}{Mn} < \ln M + 2C. \qquad (32)$$

Scegliamo ora la costante M in modo che il secondo membro della (31) sia uguale a uno:

$$\ln M - 2C = 1,$$

cioè

$$M = e^{2C+1},$$

e poniamo

$$L = M(\ln M + 2C).$$

Per il numero T dei primi compresi fra n e Mn otteniamo allora, dalle (31) e (32), le disuguaglianze

$$\frac{n}{\ln n} < T < L\frac{n}{\ln n}, \qquad (33)$$

che volevamo stabilire. Poiché $n/\ln n \to \infty$ quando n cresce indefinitamente, segue che anche $T \to \infty$.

§4. Il metodo di Vinogradov

Il metodo di Vinogradov applicato al problema di Goldbach.

Cercheremo in questo paragrafo di dare un'idea del metodo di Vinogradov nel caso particolare del problema di Goldbach: rappresentare un numero dispari come somma di tre numeri primi.

Un integrale per il numero di rappresentazioni di N come somma di tre primi. Sia N un numero dispari sufficientemente grande. Indichiamo con $I(N)$ il numero di rappresentazioni di N come somma di tre primi; in altre parole, il numero delle soluzioni dell'equazione

$$N = p_1 + p_2 + p_3 \qquad (34)$$

in numeri primi p_1, p_2, p_3, contando come distinte le terne ordinate differenti.

Il problema di Goldbach sarà risolto se si riuscirà a stabilire che $I(N) > 0$. Il metodo di Vinogradov permette di stabilire questo fatto per N sufficientemente grande e anche di trovare un'espressione approssimata per $I(N)$.

Si può scrivere $I(N)$ nella forma seguente:

$$I(N) = \sum_{p_1 \leqslant N} \sum_{p_2 \leqslant N} \sum_{p_3 \leqslant N} \int_0^1 e^{2\pi i(p_1+p_2+p_3-N)\alpha}\, d\alpha, \qquad (35)$$

dove le somme si estendono ai primi non maggiori di N. Infatti, per n intero e diverso da zero,

$$\int_0^1 e^{2\pi i n\alpha}\, d\alpha = \frac{1}{2\pi n i}\left[e^{2\pi i n\alpha}\right]_0^1 = \frac{1}{2\pi n i}(e^{2\pi n i} - e^0) = 0,$$

poiché

$$e^{2\pi i n} = \cos 2\pi n + i \sin 2\pi n = 1;$$

mentre, se $n = 0$,

$$\int_0^1 e^{2\pi i n\alpha}\, d\alpha = \int_0^1 d\alpha = 1.$$

Pertanto, ogni volta che i primi p_1, p_2 e p_3 hanno somma N, l'integrale sotto il segno di sommatoria nella (35) vale uno; quando invece $p_1 + p_2 + p_3 \neq N$, esso vale zero. È così dimostrata la validità della (35).

Poiché $e^{2\pi i a} \cdot e^{2\pi i b} = e^{2\pi i(a+b)}$ e l'integrale di una somma è uguale alla somma degli integrali dei suoi termini, dalla (35) segue

$$I(N) = \int_0^1 \Big(\sum_{p \leqslant N} e^{2\pi i \alpha p} \Big)^3 e^{-2\pi i \alpha N} \, d\alpha.$$

Introducendo la notazione

$$T_\alpha = \sum_{p \leqslant N} e^{2\pi i \alpha p} \qquad (36)$$

abbiamo quindi

$$I(N) = \int_0^1 T_\alpha^3 e^{-2\pi i \alpha N} \, d\alpha. \qquad (37)$$

Suddivisione dell'intervallo d'integrazione in intervalli principali e complementari. Sia h una quantità non intera scelta opportunamente in funzione di N, che cresca indefinitamente con N ma sia piccola rispetto a N e persino rispetto a $\sqrt[3]{N/2}$, e poniamo $\tau = N/h$. Poiché la funzione integranda nella (37) ha periodo uno, l'intervallo d'integrazione si può sostituire con il segmento da $-1/\tau$ a $1 - 1/\tau$. Dunque

$$I(N) = \int_{-1/\tau}^{1-1/\tau} T_\alpha^3 e^{-2\pi i \alpha N} \, d\alpha. \qquad (38)$$

Consideriamo ora tutte le frazioni proprie irriducibili a/q, comprendendo 0/1, con denominatore non maggiore di h, e distinguiamo nel segmento $-(1/\tau) \leqslant \alpha \leqslant 1 - (1/\tau)$ gli intervalli «principali» corrispondenti a queste frazioni:

$$\frac{a}{q} - \frac{1}{\tau} \leqslant \alpha \leqslant \frac{a}{q} + \frac{1}{\tau}; \qquad (39)$$

Per N sufficientemente grande questi intervalli, come si può dimostrare,[6] non hanno punti in comune. In tal modo il segmento $-(1/\tau) \leqslant \alpha \leqslant 1 - (1/\tau)$ si scompone in intervalli principali e intervalli «complementari».

Rappresentiamo $I(N)$ come somma di due termini:

$$I(N) = I_1(N) + I_2(N), \qquad (40)$$

dove $I_1(N)$ indica la somma degli integrali sugli intervalli principali e $I_2(N)$ la somma degli integrali sugli intervalli complementari. Come vedremo, al crescere indefinito di N attraverso valori dispari cresce indefinitamente anche $I_1(N)$, mentre

$$\lim_{N \to \infty} \frac{I_2(N)}{I_1(N)} = 0. \qquad (41)$$

Dalla (40) risulta quindi che il numero di rappresentazioni di un intero dispari N come somma di tre primi cresce indefinitamente con N; in particolare, avremo dimostrato la congettura di Goldbach per tutti gli interi dispari sufficientemente grandi.

Un'espressione per l'integrale sugli intervalli principali. Sia α appartenente a uno degli intervalli principali. Dalla (39), $\alpha = a/q + z$, dove $1 \leqslant q \leqslant h$ e $|\, z \,| \leqslant 1/\tau$. Scomponiamo la somma (36)

$$T_\alpha = \sum_{p \leqslant N} e^{2\pi i p \alpha} = \sum_{p \leqslant N} e^{2\pi i (a/q+z) p},$$

estesa a tutti i primi non maggiori di N, in somme parziali $T_{\alpha,M}$ della forma

$$T_{\alpha,M} = \sum_{M \leqslant p < M'} e^{2\pi i (a/q+z) p},$$

dove M' è scelto in modo che $e^{2\pi i z p}$ differisca «poco» da $e^{2\pi i z M}$. Poiché intendiamo dare soltanto un'idea del metodo di Vinogradov, e non una dimostrazione del teorema di Goldbach-Vinogradov, non preciseremo il significato dell'espressione «differisce poco». Nella sua dimostrazione Vinogradov utilizza disuguaglianze rigorosamente definite, che richiedono molti calcoli. Abbiamo dunque

$$T_{\alpha,M} \approx e^{2\pi i M z} \sum_{M \leqslant p < M'} e^{2\pi i (a/q) p} = e^{2\pi i M z}\, T_{a/q,M}, \qquad (42)$$

dove il simbolo $\approx$ indica che la prima delle tre espressioni dell'ultima relazione differisce «poco» dalla seconda.

Scomponiamo ulteriormente ciascuna delle somme

$$T_{a/q,M} = \sum_{M \leqslant p < M'} e^{2\pi i (a/q) p} \qquad (43)$$

in somme $T_{a/q,M,l}$, estese a tutti i primi p_l che soddisfano la relazione $M \leqslant p_l < M'$ e appartengono alle progressioni aritmetiche $qx + l$, dove l assume tutti i valori da 0 a $q - 1$ primi con q. I pochi primi che dividono q sono separati in un termine d'errore nell'approssimazione. Ma

$$e^{2\pi i (a/q) p l} = e^{2\pi i x + 2\pi i (a/q) l} = e^{2\pi i (a/q) l},$$

e quindi

$$T_{a/q,M'l} = e^{2\pi i (a/q) l} \cdot \pi(M, M', l), \qquad (44)$$

dove $\pi(M, M', l)$ è il numero dei primi che soddisfano le condizioni $M \leqslant p < M'$ e appartengono alla progressione aritmetica $qx + l$. Nello sviluppo della formula (14) per il numero $\pi(x)$ dei primi non maggiori di x, si stabilì che $\pi(M, M', l)$, per valori di q «piccoli» rispetto alla differenza $M' - M$, differisce poco da $1/\phi(q) \int_M^{M'} dx/\ln x$, dove $\phi(q)$ è la funzione di Eulero. È una funzione aritmetica, cioè definita per i numeri naturali q, che conta gli interi positivi non maggiori di q e primi con q. Dalla (44) possiamo dunque ricavare

$$T_{a/q,M'l} \approx e^{2\pi i(a/q)l} \cdot \frac{1}{\phi(q)} \int_M^{M'} \frac{dx}{\ln x}. \qquad (45)$$

Nell'espressione a destra della (45), solo il primo fattore dipende da l, cioè dalla scelta della progressione aritmetica $qx + l$; consideriamo ora q fissato. Sommando rispetto a l, otteniamo

$$T_{a/q,M} \approx \frac{1}{\phi(q)} \int_M^{M'} \frac{dx}{\ln x} \sum_l e^{2\pi i(a/q)l},$$

e quindi, dalla (42),

$$T_{\alpha,M} \approx e^{2\pi iMz} \cdot \frac{1}{\phi(q)} \int_M^{M'} \frac{dx}{\ln x} \cdot \sum_l e^{2\pi i(a/q)l}, \qquad (46)$$

dove

$$e^{2\pi iMz} \int_M^{M'} \frac{dx}{\ln x} \approx \int_M^{M'} \frac{e^{2\pi izx}}{\ln x} dx,$$

Ciò permette di sostituire la (46) con la relazione

$$T_{\alpha,M} \approx \int_M^{M'} \frac{e^{2\pi izx}}{\ln x} dx \cdot \frac{1}{\phi(q)} \sum_l e^{2\pi i(a/q)l}. \qquad (47)$$

Sommando rispetto a M, si stabilisce che

$$T_\alpha \approx \int_2^N \frac{e^{2\pi izx}}{\ln x} dx \cdot \frac{1}{\phi(q)} \sum_l e^{2\pi i(a/q)l}. \qquad (48)$$

La somma

$$\sum_l e^{2\pi i(a/q)l},$$

che compare a destra nella (48), estesa ai naturali l non maggiori di q e primi con q, si può esprimere mediante una funzione aritmetica $\mu(q)$ così definita: $\mu(q) = 0$ se q è divisibile per il quadrato di un intero maggiore di uno; $\mu(1) = 1$; infine, $\mu(q) = (-1)^n$ se $q = p_1p_2...p_n$, dove p_1, p_2, ..., p_n sono primi distinti. Per a e q primi fra loro, dunque,

$$\sum_l e^{2\pi i(a/q)l} = \mu(q). \qquad (49)$$

La (48) si può dunque scrivere nella forma

$$T_\alpha \approx \frac{\mu(q)}{\phi(q)} \int_2^N \frac{e^{2\pi izx}}{\ln x} dx.$$

Poiché $\mu^3(q) = \mu(q)$, abbiamo

$$T_\alpha^3 \approx \frac{\mu(q)}{(\phi(q))^3}\left(\int_2^N \frac{e^{2\pi i z x}}{\ln x}\,dx\right)^3, \tag{50}$$

e dalla definizione di $I_1(N)$

$$I_1(N) = \sum_{1\leqslant q<h}\sum_a \int_{a/q-1/\tau}^{a/q+1/\tau} T_\alpha^3\, e^{-2\pi i \alpha N}\,d\alpha, \tag{51}$$

dove, per ogni q fissato, la somma è estesa agli interi a non negativi minori di q e primi con q. Poiché $\alpha = a/q + z$, dalla (50) otteniamo

$$I_1(N) \approx \sum_{1\leqslant q<h} \frac{\mu(q)}{(\phi(q))^3}\sum_a e^{-2\pi i (a/q)N}\int_{-1/\tau}^{1/\tau}\left(\int_2^N \frac{e^{2\pi i z x}}{\ln x}\,dx\right)^3 e^{-2\pi i z N}\,dz. \tag{52}$$

Introduciamo la notazione

$$R(N) = \int_{-1/\tau}^{1/\tau}\left(\int_2^N \frac{e^{2\pi i z x}}{\ln x}\,dx\right)^3 e^{-2\pi i z N}\,dz. \tag{53}$$

Dalla relazione (52) segue

$$I_1(N) \approx R(N)\sum_{1\leqslant q<h}\frac{\mu(q)}{[\phi(q)]^3}\sum_a e^{-2\pi i (a/q)N}. \tag{54}$$

Occorre qui richiamare l'attenzione sul fatto che $R(N)$ è un'espressione analitica, e quindi si può calcolare approssimativamente. Risulta infatti che

$$R(N) \approx \frac{N^2}{2(\ln N)^3}. \tag{55}$$

L'espressione che moltiplica $R(N)$ a destra nella (54) differisce «poco» dalla somma della serie infinita

$$S(N) = \sum_{q=1}^{\infty}\frac{\mu(q)}{[\phi(q)]^3}\sum_a e^{-2\pi i (a/q)N}, \tag{56}$$

Dalle (54) e (55) si può quindi stabilire che

$$I_1(N) \approx \frac{N^2}{2(\ln N)^3}S(N), \tag{57}$$

o, più precisamente,

$$I_1(N) = \frac{N^2}{2(\ln N)^3}[S(N) + \gamma_1(N)], \tag{58}$$

dove

$$\lim_{N\to\infty} \gamma_1(N) = 0. \qquad (59)$$

Osserviamo che l'espressione aritmetica $S(N)$ si può scrivere nella forma

$$S(N) = C \prod_p \left(1 - \frac{1}{p^2 - 3p + 3}\right), \qquad (60)$$

dove C è una costante, il prodotto si estende a tutti i divisori primi del numero N e, come mostrano i calcoli,

$$S(N) > 0.6. \qquad (61)$$

Stima dell'integrale sugli intervalli complementari. Passiamo ora a stimare la somma I_2 degli integrali sugli intervalli complementari. Poiché il modulo di un integrale non supera l'integrale del modulo della funzione integranda e $|e^{-2\pi i\alpha N}| = 1$ per αN reale, abbiamo

$$| I_2 | < \max | T_\alpha | \cdot \int_{-1/\tau}^{1-1/\tau} | T_\alpha |^2 \, d\alpha, \qquad (62)$$

dove max $|T_\alpha|$ rappresenta il massimo valore di $|T_\alpha|$ per α appartenente agli intervalli complementari. Abbiamo ulteriormente maggiorato il secondo membro prendendo, come fattore di max $|T_\alpha|$, l'integrale esteso a tutto l'intervallo $-(1/\tau) \leqslant \alpha \leqslant 1 - 1/\tau$.

Ma il quadrato del modulo di un numero complesso è uguale al prodotto del numero per il suo coniugato; dunque

$$| T_\alpha |^2 = T_\alpha \cdot \overline{T}_\alpha ,$$

dove, dalla (36), abbiamo

$$\overline{T}_\alpha = \sum_{p \leqslant N} e^{-2\pi i\alpha p},$$

poiché $e^{-2\pi i\alpha p} = \cos 2\pi\alpha p - i \sin 2\pi\alpha p$. La disuguaglianza (62) si può pertanto riscrivere nella forma

$$| I_2 | < \max | T_\alpha | \cdot \int_{-1/\tau}^{1-1/\tau} \sum_{p \leqslant N} e^{2\pi i\alpha p} \sum_{p_1 \leqslant N} e^{-2\pi i\alpha p_1} \, d\alpha$$

oppure nella forma

$$| I_2 | < \max | T_\alpha | \cdot \int_{-1/\tau}^{1-1/\tau} \sum_{p \leqslant N} \sum_{p_1 \leqslant N} e^{2\pi i\alpha(p - p_1)} \, d\alpha. \qquad (63)$$

In base a quanto detto all'inizio del paragrafo, l'integrale nella disuguaglianza (63) rappresenta il numero U delle soluzioni dell'equazione $p - p_1 = 0$ in primi p, p_1 non maggiori

di N; è dunque semplicemente il numero dei primi non maggiori di N, cioè $\pi(N)$. Dal risultato (12) di Čebyšev abbiamo

$$\pi(N) < B \cdot \frac{N}{\ln N},$$

dove B è una costante. In tal modo

$$|I_2| < B \cdot \frac{N}{\ln N} \cdot \max |T_\alpha|, \qquad (64)$$

dove, ripetiamo, max $|T_\alpha|$ è il massimo valore di $|T_\alpha|$ sugli intervalli complementari. Dalle (58) e (59) segue che, per completare la dimostrazione del teorema di Goldbach-Vinogradov, dobbiamo ora mostrare che max $|T_\alpha|$ è di ordine inferiore a $N/(\ln N)^2$. Stabilire questo fatto presenta però la difficoltà maggiore e costituisce la parte essenziale dell'intera dimostrazione.

Ogni α appartenente a un intervallo complementare si può rappresentare nella forma $\alpha = a/q + z$, dove $h < q \leqslant \tau$ e $|z| \leqslant 1/q\tau$. Il problema consiste dunque nello stimare il modulo della somma trigonometrica

$$T_\alpha = \sum_{p \leqslant N} e^{2\pi i(a/q+z)p}$$

sotto le condizioni assegnate. Vinogradov stabilì, in particolare, che

$$\lim_{N\to\infty} \max |T_\alpha| / [N/(\ln N)^3] = 0. \qquad (65)$$

Egli utilizzò a questo scopo un'identità importantissima, da lui scoperta, per la funzione $\mu(n)$ già discussa.

Purtroppo non è possibile dare qui una dimostrazione della (65); rimandiamo il lettore interessato al capitolo X del libro di Vinogradov «Il metodo delle somme trigonometriche nella teoria dei numeri».

Dalle (65) e (64), come si è osservato, segue

$$\lim_{N\to\infty} \frac{I_2(N)}{I_1(N)} = 0.$$

In tal modo, dalle (40), (58) e (59), abbiamo

$$I(N) = \frac{N^2}{2(\ln N)^3}[S(N) + \gamma(N)], \qquad (66)$$

dove

$$\lim_{N\to\infty} \gamma(N) = 0,$$

e $S(N)$ ha il valore (60), cosicché, dalla (61), $S(N) > 0{,}6$. Si conclude così la dimostrazione del teorema, una volta ammesse le stime qui enunciate.

§5. Somme di due quadrati e interi complessi

L'importanza dello studio dei numeri primi dipende soprattutto dal ruolo centrale che essi svolgono nella maggior parte delle leggi della teoria dei numeri. Accade spesso che questioni apparentemente lontane dalla divisibilità, esaminate più attentamente, risultino intimamente connesse con la teoria dei primi. Illustriamo questa affermazione con un esempio.

Uno dei problemi della teoria dei numeri consiste nel trovare quali numeri naturali si possano scomporre nella somma dei quadrati di due interi, non necessariamente diversi da zero.

La regola che governa la successione dei numeri rappresentabili come somma di due quadrati non è subito chiara. Da 1 a 50, per esempio, troviamo 1, 2, 4, 5, 8, 9, 10, 13, 16, 18, 20, 25, 26, 29, 32, 34, 36, 37, 40, 41, 45, 49, 50: una successione che sembra piuttosto irregolare. Il matematico francese del XVII secolo Fermat osservò che tutto dipende da come il numero si rappresenta come prodotto di primi: la questione è dunque intrinsecamente legata alla teoria dei numeri primi.

I primi diversi da $p = 2$ sono dispari; dividendo per 4, danno pertanto resto 1, se sono della forma $4n + 1$, oppure resto 3, se sono della forma $4n + 3$.

Esamineremo la questione della rappresentazione di un numero come somma di due quadrati nei tre punti seguenti.

1. Un primo dispari p è somma di due quadrati se e solo se $p = 4n + 1$. Il primo 2 è anch'esso somma di due quadrati: $2 = 1^2 + 1^2$.

È quasi immediato dimostrare che un numero della forma $4n + 3$ non può essere somma di due quadrati. La somma dei quadrati di due numeri pari è divisibile per 4; quella dei quadrati di due numeri dispari dà resto 2 nella divisione per 4; quella del quadrato di un pari e del quadrato di un dispari dà resto 1.

Dimostriamo ora un teorema preliminare: se p è primo, $(p - 1)! + 1$ è divisibile per p. I numeri non divisibili per p, divisi per p, danno i resti 1, 2, 3, …, $p - 1$. Scegliamo un intero r, $1 \leqslant r \leqslant p - 1$, e moltiplichiamolo per 1, 2, …, $p - 1$. Dividendo per p i prodotti così ottenuti, si ritrovano, come non è difficile dimostrare, tutti gli stessi resti, generalmente in un ordine diverso. Fra essi vi è in particolare il numero 1: per ogni r si può quindi trovare un r_1 tale che $r \cdot r_1 = 1 + kp$. Osserviamo che $r = r_1$ soltanto se $r = 1$ oppure $r = p - 1$. Infatti, se $r^2 = 1 + kp$, allora $(r + 1)(r - 1)$ è divisibile per p; ma per i numeri $1 \leqslant r \leqslant p - 1$ ciò è possibile soltanto per $r = 1$ e $r = p - 1$. Troviamo il resto della divisione di $(p - 1)! = 1 \cdot 2 \cdot \ldots \cdot (p - 1)$ per p. In questo prodotto, a ogni fattore r, esclusi 1 e $p - 1$, corrisponde un fattore r_1 distinto da r, tale che $r \cdot r_1$ dia resto 1. Per p dispari, dunque, il resto di $(p - 1)!$ è lo stesso che si otterrebbe conservando soltanto i due fattori 1 e $p - 1$: è cioè $p - 1$. Pertanto $(p - 1)! + 1$ è divisibile per p. Per $p = 2$ il risultato si verifica immediatamente.

Sia ora $p = 4n + 1$. Scriviamo

$$(p-1)! + 1 = \left\{1 \cdot 2 \cdots \frac{p-1}{2}\right\} \cdot \left\{\left(p - \frac{p-1}{2}\right) \cdots (p-2)(p-1)\right\} + 1.$$

Il secondo fattore fra parentesi graffe, diviso per p, ha lo stesso resto di $(-1)^{(p-1)/2}\,[(p-1)/2]!$. Ma $(p-1)/2 = 2n$ è pari, e quindi anche $\{[(p-1)/2]!\}^2 + 1$ è divisibile per p. Indichiamo con A il resto della divisione di $[(p-1)/2]!$ per p. Evidentemente anche $A^2 + 1$ è divisibile per p.

Consideriamo l'espressione $x - Ay$, in cui x e y percorrono indipendentemente i numeri $0, 1, \cdots, [\sqrt{p}]$; qui $[x]$ indica il massimo intero non maggiore di x. Otteniamo così $([\sqrt{p}] + 1)^2 \geqslant p + 1$ valori numerici di $x - Ay$, che possono essere distinti oppure, in alcuni casi, coincidere. I resti possibili nella divisione per p sono soltanto p, cioè 0, 1, 2, ..., $p - 1$, mentre abbiamo almeno $p + 1$ valori di $x - Ay$. Devono quindi esistere due coppie distinte (x_1, y_1) e (x_2, y_2) per cui $x_1 - Ay_1$ e $x_2 - Ay_2$ diano lo stesso resto. In altre parole, $(x_1 - x_2) - A(y_1 - y_2)$ è divisibile per p. Poniamo $x_0 = x_1 - x_2$, $y_0 = y_1 - y_2$. Evidentemente $|x_0| < \sqrt{p}, |y_0| < \sqrt{p}$. Poiché $A^2 + 1$ è divisibile per p, anche $y_0^2(A^2 + 1) = (Ay_0)^2 + y_0^2$ è divisibile per p; inoltre, poiché $x_0 - Ay_0$ è divisibile per p, lo è anche il numero $x_0^2 - (Ay_0)^2 = (x_0 - Ay_0)(x_0 + Ay_0)$. Dunque la quantità $x_0^2 + y_0^2$, uguale a $(x_0^2 - (Ay_0)^2 + (Ay_0)^2 + y_0^2)$, è divisibile per p. Ma $|x_0| < \sqrt{p}, |y_0| < \sqrt{p}$. Ne segue $x_0^2 + y_0^2 = 0$ oppure $x_0^2 + y_0^2 = p$. Il primo caso è impossibile, perché le coppie (x_1, y_1) e (x_2, y_2) sono distinte. Pertanto un primo della forma $4n + 1$ si può rappresentare come somma di due quadrati.

2. Passiamo alla scomposizione di un intero positivo qualsiasi nella somma di due quadrati. È facile stabilire l'identità

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

Questa identità mostra che il prodotto di due interi, ciascuno somma di due quadrati, è ancora somma di due quadrati. Pertanto il prodotto di potenze qualsiasi di primi della forma $4n + 1$, oppure uguali a 2, è somma di due quadrati. Poiché moltiplicando una somma di due quadrati per un quadrato si ottiene ancora una somma di due quadrati, ogni numero nella cui fattorizzazione i primi della forma $4n + 3$ compaiono con esponenti pari è somma di due quadrati.

3. Mostriamo ora che, se un primo della forma $4n + 3$ compare nella fattorizzazione di un numero con esponente dispari, quel numero non si può esprimere come somma di due quadrati. La questione iniziale sarà così completamente risolta.

Consideriamo i numeri complessi della forma $a + bi$, dove a e b sono interi ordinari. Un tale numero si chiamerà intero complesso. Se un intero N è somma di due quadrati, $N = a^2 + b^2$, allora $N = (a + bi)(a - bi) = \alpha \cdot \bar{\alpha}$, dove $\bar{\alpha}$ indica il coniugato del numero α. In altre parole, nel dominio degli interi complessi N si scompone in fattori complessi coniugati.

In questo dominio si può costruire una teoria della divisibilità del tutto analoga a quella degli interi ordinari. Diremo che l'intero complesso α è divisibile per l'intero complesso non nullo β se α/β è ancora un intero complesso. Esistono soltanto quattro interi complessi che dividono

1, detti unità: 1, −1, i e $-i$. Diremo primo un intero complesso α non nullo e non unità che non abbia divisori diversi da 1, −1, i, $-i$, α, $-\alpha$, αi, $-\alpha i$. Il problema risolto nel primo punto assume ora un altro significato: i numeri della forma $4n + 1$, e il numero 2, che prima erano primi, cessano di esserlo nel dominio degli interi complessi; è invece facile dimostrare che i primi della forma $4n + 3$ rimangono primi.

Infatti, se $p = \alpha\beta$, allora $p = \bar{\alpha}\bar{\beta}$ e $p^2 = \alpha\bar{\alpha}\beta\bar{\beta}$. Ma $\alpha\bar{\alpha}$ e $\beta\bar{\beta}$ sono interi positivi ordinari; inoltre $p \neq \alpha\bar{\alpha}$, perché i primi della forma $4n + 3$ non sono somma di due quadrati. Il prodotto delle due norme è p^2, e nessuna delle due può essere p: una è dunque 1 e l'altra p^2. Scambiando se necessario α e β, abbiamo $\alpha\bar{\alpha} = 1$. Quindi α può essere soltanto ±1 oppure $\pm i$: p non ha divisori diversi da quelli ovvi.

Per gli interi complessi vale ancora il teorema dell'unicità della scomposizione in fattori primi. L'unicità si intende naturalmente a meno dell'ordine dei fattori e della moltiplicazione per le unità 1, −1, i, $-i$.

Sia N somma di due quadrati, $N = \alpha\bar{\alpha}$, e sia p un primo della forma $4n + 3$. Calcoliamo con quale esponente p compare in N. Poiché p rimane primo nel dominio complesso, basta determinare gli esponenti con cui compare in α e in $\bar{\alpha}$. Ma questi esponenti sono uguali, cosicché p compare necessariamente in N con esponente pari. La proposizione è dimostrata.

La scoperta che una ricca teoria della divisibilità è possibile anche fuori dal dominio degli interi razionali ampliò enormemente l'orizzonte dei matematici del XIX secolo. Lo sviluppo di queste idee richiese la creazione di nuovi concetti generali, per esempio quelli di anello e di ideale. Il significato di tali concetti ha ormai oltrepassato di molto i confini della teoria dei numeri.

Letture consigliate

H. Davenport, The higher arithmetic: an introduction to the theory of numbers, Harper Brothers, New York, 1960.

T. Estermann, Introduction to modern prime number theory, Cambridge University Press, New York, 1952.

A. A. Fraenkel, Integers and theory of numbers, Scripta Mathematica, New York, 1955.

G. H. Hardy and F. M. Wright, An introduction to the theory of numbers, 2nd ed., Clarendon Press, Oxford, 1945.

A. Ia. Khinchin, Three pearls of number theory, Graylock Press, Rochester, N. Y., 1952.

O. Ore, Number theory and its history, McGraw-Hill, New York, 1948.

I. M. Vinogradov, An introduction to the theory of numbers, Pergamon Press, New York, 1955.

Note al capitolo X

[1] Si veda il capitolo IV.

[2] Abbiamo qui un esempio di equazione indeterminata, da studiare dal punto di vista della sua risolubilità in numeri interi.

[3] Se $a_0 = 1$, il numero algebrico si chiama intero algebrico. Un numero che non è algebrico si chiama trascendente.

[4] Nel 1949 A. Selberg diede una dimostrazione elementare, cioè senza ricorrere a variabili complesse, della legge asintotica della distribuzione dei primi.

[5] La validità della congettura secondo cui ogni numero pari *N* sufficientemente grande si può rappresentare come somma di due primi rimane una questione aperta all'epoca dell'opera.

[6] Se due di questi intervalli, centrati nei punti a_1/q_1 e a_2/q_2, si intersecano, in un punto comune avremo l'uguaglianza

$$a_1/q_1 + \theta_1/\tau = a_2/q_2 + \theta_2/\tau,$$
$$\text{dove } |\theta_1| \le 1, |\theta_2| \le 1,$$

ossia

$$(a_1q_2 - a_2q_1)/(q_1q_2) = (\theta_2 - \theta_1)/\tau.$$

Ma il valore assoluto del primo membro di quest'ultima equazione non è minore di $1/(q_1q_2)$, e quindi è almeno $1/h^2$; quello del secondo membro non supera $2/\tau = 2h/N$. Se l'uguaglianza fosse vera, ne seguirebbe $1/h^2 \le 2h/N$, in contraddizione con la scelta di h.

Nota editoriale

La traduzione conserva le cinque sezioni, la figura, le formule, la bibliografia e le sei note del capitolo. Le parole inglesi presenti in alcune formule sono tradotte. Sono corretti refusi in indici, esponenti e rinvii: il prodotto nella prova di Euclide contiene $p_1p_2…p_k$; il primo membro della (19) è ln(n!); nell'approssimazione razionale il limite dell'errore è 1/(qτ); nella (65) si considera il modulo della somma. Nella prova sui due quadrati è ripristinato il fattore $x_0 + Ay_0$ nella differenza di quadrati. Nella nota 6 è corretto il segno della differenza $\theta_2 - \theta_1$.

Si rendono espliciti il coefficiente direttivo non nullo nella definizione di numero algebrico, il polo di ζ(s) in s = 1, il significato del numero di termini nella (17), la natura ordinata delle terne contate da I(N) e la condizione di coprimalità nelle somme sui residui. Per uniformare i limiti q < h e q ≤ h si può scegliere h non intero. Nel metodo di Vinogradov i primi che dividono il modulo sono assorbiti nei termini d'errore; il testo fornisce uno schema della dimostrazione, fondato sulle stime enunciate. Per il risultato di Linnik si usa un modulo k > 1 e una classe residua prima con k.

Il crivello include i primi fino alla radice quadrata del limite assegnato. Il criterio p = 4n + 1 riguarda i primi dispari: 2 è trattato separatamente. Nella definizione di primo complesso sono esclusi zero e le unità; nella prova per i primi 4n + 3 si esplicita lo scambio dei fattori necessario per ottenere una norma uguale a uno. Le approssimazioni di π sono indicate con ≈. I riferimenti allo stato delle ricerche, alle congetture aperte e ai migliori risultati disponibili conservano il significato storico dell'opera e non costituiscono un aggiornamento contemporaneo.

CAPITOLO XI

La teoria della probabilità

A. N. Kolmogorov

§1. Le leggi della probabilità

Le leggi più semplici delle scienze naturali sono quelle che stabiliscono le condizioni nelle quali un evento che ci interessa si verifica certamente oppure certamente non si verifica. Esse si possono esprimere in una delle due forme seguenti:

1. Se si realizza un complesso, cioè un insieme, di condizioni *S*, l'evento *A* si verifica certamente.

2. Se si realizza un complesso di condizioni *S*, l'evento *A* non può verificarsi.

Nel primo caso l'evento *A*, rispetto al complesso di condizioni *S*, si chiama «certo» o «necessario»; nel secondo, «impossibile». Per esempio, alla pressione atmosferica e a una temperatura *t* compresa fra 0 °C e 100 °C, che costituiscono il complesso di condizioni *S*, l'acqua pura in equilibrio si trova allo stato liquido: l'evento A_1 è certo. Non si trova invece interamente allo stato gassoso o solido: gli eventi A_2 e A_3, così intesi, sono impossibili.

Un evento *A* che, sotto un complesso di condizioni *S*, talvolta si verifica e talvolta no si chiama casuale rispetto a quel complesso di condizioni. Sorge allora una domanda: la casualità di *A* dimostra forse che non esiste alcuna legge che colleghi le condizioni *S* all'evento *A*? Supponiamo, per esempio, che lampade di un certo tipo prodotte in una determinata fabbrica, condizione *S*, continuino talvolta a funzionare per oltre 2 000 ore, evento *A*, mentre altre si brucino e diventino inutilizzabili prima. Non potrebbero i risultati delle prove sulla durata delle lampade servire ugualmente a valutare la produzione della fabbrica? Oppure dovremmo limitarci a indicare un periodo, diciamo 500 ore, durante il quale praticamente tutte le lampade funzionano, e un altro, diciamo 10 000 ore, dopo il quale praticamente nessuna funziona più? È chiaro che descrivere la durata di una lampada mediante una disuguaglianza della forma $500 \leqslant T \leqslant 10{,}000$ aiuta poco il consumatore. Egli riceverà un'informazione assai più utile sapendo che circa l'80% delle lampade funziona per almeno 2 000 ore. Una valutazione ancora più completa della qualità consiste nell'indicare, per ogni *T*, la percentuale $v(T)$ delle lampade che funzionano per almeno *T* ore, per esempio mediante il grafico della figura 1.

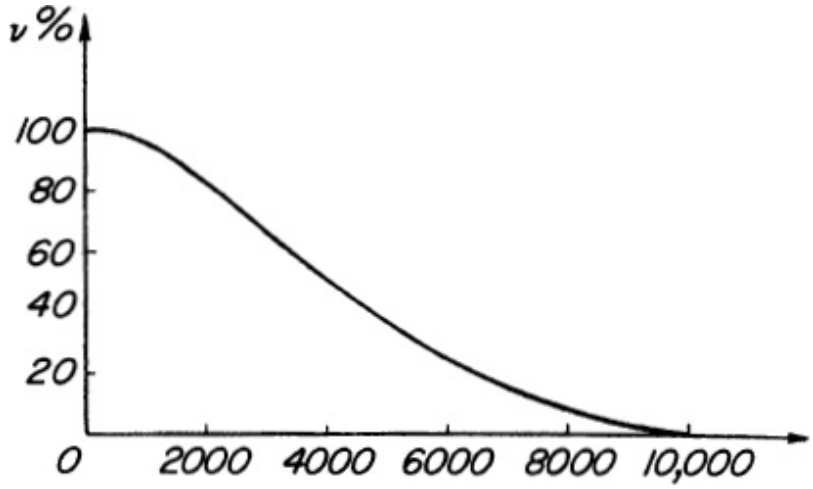


Figura 1.

In pratica, la curva $v(T)$ si ricava provando un campione sufficientemente numeroso di lampade, nell'esempio 100-200. La curva così ottenuta ha valore reale solo se rappresenta una legge che riguarda sia il campione esaminato sia tutte le lampade fabbricate con materiali della stessa qualità e nelle stesse condizioni tecnologiche. Occorre cioè che le medesime prove, condotte su un altro campione, diano risultati approssimativamente uguali: la nuova curva $v(T)$ deve differire poco dalla prima. In altre parole, la legge statistica espressa dalle curve $v(T)$ dei diversi campioni riflette una legge probabilistica che collega la durata utile della lampada ai materiali e alle condizioni tecnologiche della sua fabbricazione.

Questa legge probabilistica è data da una funzione $P(T)$, dove $P(T)$ è la probabilità che una singola lampada, prodotta nelle condizioni assegnate, funzioni per almeno T ore.

Affermare che l'evento A si verifica, sotto le condizioni S, con una probabilità determinata

$$\mathbf{P}(A/S) = p$$

equivale a dire che, in serie sufficientemente lunghe di prove, cioè di realizzazioni del complesso di condizioni S, le frequenze

$$\nu_r = \frac{\mu_r}{n_r}$$

dell'evento A saranno approssimativamente uguali fra loro e vicine a p. Qui n_r è il numero delle prove nella r-esima serie e μ_r il numero di quelle in cui si verifica A.

L'ipotesi che esista una costante $p = P(A/S)$, determinata oggettivamente dal legame fra le condizioni S e l'evento A, alla quale le frequenze v si avvicinano «in generale» al crescere del numero delle prove, trova un buon riscontro pratico per una vasta classe di eventi. Gli eventi di questo tipo sono abitualmente detti casuali o stocastici.

L'esempio precedente riguarda le leggi probabilistiche della produzione di massa. La realtà di queste leggi è indubbia, ed esse sono alla base di importanti applicazioni pratiche nel controllo statistico della qualità. Analoghe sono le leggi probabilistiche della dispersione dei

proiettili, fondamentali nella teoria del tiro. Poiché storicamente questa fu una delle prime applicazioni della probabilità a problemi tecnici, torneremo più avanti su alcuni semplici problemi di tale teoria.

Quanto abbiamo detto sulla «vicinanza» della frequenza v alla probabilità p, per un grande numero n di prove, è piuttosto vago: non abbiamo precisato quanto piccola possa essere la differenza $v - p$ per un dato n. Il grado di vicinanza sarà stimato nel §3. È interessante osservare che una certa indeterminatezza è inevitabile: l'affermazione stessa che v e p siano vicini ha soltanto carattere probabilistico, come risulta quando si cerca di precisare l'intera questione.

§2. Assiomi e formule fondamentali della teoria elementare della probabilità

Poiché l'importanza delle leggi statistiche è indubbia, passiamo ai metodi per studiarle. Si pensa anzitutto alla possibilità di procedere in modo puramente empirico. Dato che una legge probabilistica si manifesta soltanto nei fenomeni di massa, sembra naturale supporre che, per scoprirla, occorra compiere un esperimento di massa.

Questa idea è però giusta soltanto in parte. Una volta stabilite sperimentalmente alcune leggi probabilistiche, possiamo dedurne altre con il ragionamento logico o con il calcolo, sotto determinate ipotesi generali. Prima di mostrare come, dobbiamo elencare alcune definizioni e formule fondamentali della teoria della probabilità.

Dall'interpretazione della probabilità come valore tipico della frequenza $v = m/n$, dove $0 \leqslant m \leqslant n$ e quindi $0 \leqslant \nu \leqslant 1$, segue che la probabilità P(A) di un evento qualsiasi A va assunta compresa fra zero e uno:[1]

$$0 \leqslant \mathbf{P}(A) \leqslant 1. \qquad (1)$$

Due eventi si dicono incompatibili se non possono verificarsi entrambi sotto il complesso di condizioni S. Per esempio, nel lancio di un dado, ottenere un numero pari di punti e ottenere un tre sono eventi incompatibili. Un evento A si chiama unione degli eventi A_1 e A_2 se consiste nel verificarsi di almeno uno di essi. Così, nel lancio di un dado, l'evento A «ottenere 1, 2 o 3» è l'unione di A_1 «ottenere 1 o 2» e A_2 «ottenere 2 o 3». Si vede facilmente che, per i numeri di occorrenze m_1, m_2 e m di due eventi incompatibili A_1, A_2 e della loro unione $A = A_1 \cup A_2$, vale $m = m_1 + m_2$; per le frequenze corrispondenti, $v = v_1 + v_2$.

Si giunge così naturalmente al seguente assioma di addizione delle probabilità:

$$\mathbf{P}(A_1 \cup A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2), \qquad (2)$$

se A_1 e A_2 sono incompatibili e $A_1 \cup A_2$ indica la loro unione.

Inoltre, per un evento certo U, poniamo naturalmente

$$\mathbf{P}(U) = 1. \qquad (3)$$

L'intera teoria matematica della probabilità si costruisce sulla base di assiomi semplici del tipo (1), (2) e (3). Dal punto di vista della matematica pura, la probabilità è una funzione numerica degli «eventi», dotata di proprietà stabilite mediante assiomi. Le proprietà espresse dalle (1), (2) e (3) sono sufficienti per costruire la cosiddetta teoria elementare della probabilità, se non si pretende di includere nell'assiomatizzazione anche i concetti di evento, di unione e di intersezione, che definiremo più avanti. Per chi comincia è più utile limitarsi a una comprensione intuitiva dei termini «evento» e «probabilità», rendendosi però conto che il loro significato nella vita pratica non può essere formalizzato completamente. Ciò non compromette la piena precisione formale di una presentazione assiomatica, puramente matematica, della teoria.

L'unione di un numero qualsiasi di eventi $A_1, A_2, \ldots, A_s$ è l'evento A che consiste nel verificarsi di almeno uno di essi. Dall'assioma di addizione si ottiene facilmente, per un numero qualsiasi di eventi a due a due incompatibili e per la loro unione A,

$$\mathbf{P}(A) = \mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_s)$$

il cosiddetto teorema di addizione delle probabilità.

Se l'unione di questi eventi è un evento certo, cioè se sotto le condizioni S deve verificarsi uno degli eventi A_k, allora

$$\mathbf{P}(A_1) + \mathbf{P}(A_2) + \cdots + \mathbf{P}(A_s) = 1.$$

In questo caso gli eventi $A_1, \ldots, A_s$, a due a due incompatibili, formano un sistema completo di eventi.

Consideriamo ora due eventi A e B che, in generale, non siano incompatibili. L'evento C si chiama intersezione di A e B, e si scrive $C = AB$, se consiste nel verificarsi sia di A sia di B.[2]

Per esempio, se A è «ottenere un numero pari nel lancio di un dado» e B è «ottenere un multiplo di tre», allora C è «ottenere un sei».

In un grande numero n di prove ripetute, supponiamo che A si verifichi m volte e B si verifichi l volte, delle quali k insieme ad A. Il quoziente k/m, per $m > 0$, si chiama frequenza condizionata di B dato A. Le frequenze k/m, m/n e k/n sono legate dalla formula

$$\frac{k}{m} = \frac{k}{n} : \frac{m}{n}$$

che suggerisce naturalmente la definizione seguente.

La probabilità condizionata $\mathrm{P}(B/A)$ dell'evento B dato l'evento A è il quoziente

$$\mathbf{P}(B/A) = \frac{\mathbf{P}(AB)}{\mathbf{P}(A)}.$$

Qui si suppone, naturalmente, che $P(A) \neq 0$.

Se gli eventi A e B non hanno alcun legame essenziale, è naturale supporre che B non si verifichi più spesso, né meno spesso, quando si verifica A rispetto a quando A non si verifica. Avremo quindi, approssimativamente, $k/m \sim l/n$, ossia

$$\frac{k}{n} = \frac{k}{m}\frac{m}{n} \sim \frac{l}{n}\frac{m}{n}.$$

In quest'ultima uguaglianza approssimata, $m/n = \nu_A$ è la frequenza di A, $l/n = \nu_B$ quella di B, e $k/n = \nu_{AB}$ quella della loro intersezione.

Vediamo che queste frequenze sono legate dalla relazione

$$\nu_{AB} \sim \nu_A \nu_B .$$

Per le probabilità degli eventi A, B e AB è dunque naturale assumere la corrispondente uguaglianza esatta

$$\mathbf{P}(AB) = \mathbf{P}(A) \cdot \mathbf{P}(B). \qquad (4)$$

L'equazione (4) serve a definire l'indipendenza dei due eventi A e B.

Analogamente si può definire l'indipendenza di un numero qualsiasi di eventi. Si può anche definire l'indipendenza di un numero qualsiasi di esperimenti: in termini approssimativi, essa significa che gli esiti di una parte degli esperimenti non dipendono da quelli dei rimanenti.[3]

Calcoliamo ora la probabilità P_k che un evento A si verifichi esattamente k volte in n prove indipendenti, in ciascuna delle quali la sua probabilità p sia la stessa. Indichiamo con $\bar{A}$ l'evento che A non si verifichi. Evidentemente,

$$\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A) = 1 - p.$$

Dalla definizione di indipendenza degli esperimenti si vede facilmente che la probabilità di una successione specificata, composta da k occorrenze di A e da $n - k$ mancate occorrenze, è uguale a

$$p^k(1 - p)^{n-k}. \qquad (5)$$

Così, per $n = 5$ e $k = 2$, la probabilità della successione $A\bar{A}A\bar{A}\bar{A}$ è $p(1 - p)p(1 - p)(1 - p) = p^2(1 - p)^3$.

Per il teorema di addizione, P_k è la somma delle probabilità di tutte le successioni con k occorrenze e $n - k$ mancate occorrenze di A. In virtù della (5), è dunque il prodotto del numero di tali successioni per $p^k(1 - p)^{n-k}$. Il numero delle successioni è evidentemente uguale al numero delle combinazioni di n elementi presi k alla volta, poiché i k esiti positivi possono occupare k posizioni qualsiasi fra le n prove.

Otteniamo infine

$$P_k = C_n^k p^k (1-p)^{n-k} \quad (k = 0, 1, 2, \cdots, n) \qquad (6)$$

la cosiddetta distribuzione binomiale.

Per vedere come si applicano definizioni e formule, consideriamo un esempio tratto dalla teoria del tiro.

Supponiamo che cinque colpi a segno siano sufficienti per distruggere un bersaglio. Ci interessa sapere se sia lecito ritenere che 40 tiri assicurino i cinque colpi necessari. Una soluzione puramente empirica procederebbe così: per un bersaglio di dimensioni date e a una distanza assegnata, si compie un grande numero di serie di tiro, diciamo 200, ciascuna di 40 colpi, e si conta quante producano almeno cinque centri. Se il risultato si ottiene, per esempio, in 195 serie su 200, la probabilità *P* è approssimativamente

$$P = \frac{195}{200} = 0.975.$$

Procedendo in questo modo puramente empirico, consumeremmo 8 000 proiettili per risolvere un semplice problema particolare. In pratica, naturalmente, non si procede così. Si comincia invece supponendo che la dispersione dei proiettili a una data distanza non dipenda dalle dimensioni del bersaglio. Gli scarti longitudinali e laterali rispetto al punto medio d'impatto seguono una legge che, per le frequenze degli scarti di diversa grandezza, è illustrata nella figura 2.

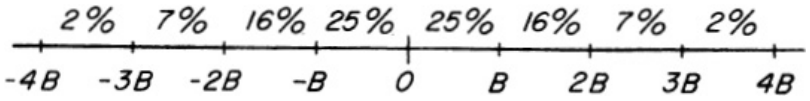


Figura 2.

La lettera *B* indica qui il cosiddetto scarto probabile. Esso è generalmente diverso nella direzione longitudinale e in quella laterale, e aumenta con la distanza. Per ogni tipo di cannone e di proiettile, gli scarti probabili alle diverse distanze si determinano empiricamente con prove di tiro in poligono. La successiva soluzione dei diversi problemi particolari del tipo descritto avviene però mediante il solo calcolo.

Per semplicità supponiamo che il bersaglio sia un rettangolo: un lato, diretto lungo la linea di tiro, è lungo quanto due scarti probabili longitudinali; l'altro, perpendicolare alla linea di tiro, è lungo quanto due scarti probabili laterali. Supponiamo inoltre che il tiro sia già ben aggiustato, così che la traiettoria media passi per il centro del bersaglio, come nella figura 3.

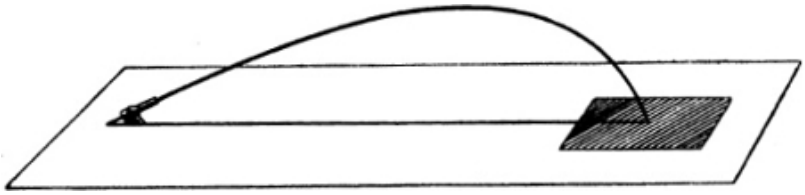


Figura 3.

Supponiamo anche che gli scarti laterali e longitudinali siano indipendenti.[4] Affinché un dato proiettile cada sul bersaglio, è allora necessario e sufficiente che i due scarti, in valore assoluto, non superino i rispettivi scarti probabili. Dalla figura 2 ciascuno di questi eventi si osserva per circa il 50% dei colpi, cioè con probabilità $\frac{1}{2}$. La loro intersezione si verifica per circa il 25% dei colpi: la probabilità che un proiettile colpisca il bersaglio è quindi

$$p = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4},$$

mentre la probabilità di mancarlo con un singolo colpo è

$$q = 1 - p = 1 - \frac{1}{4} = \frac{3}{4}.$$

Supponendo indipendenti gli esiti dei singoli colpi e applicando la formula binomiale (6), troviamo che la probabilità di ottenere esattamente k centri su 40 tiri è

$$P_k = C_{40}^{\,k} p^k q^{40-k} = [40!/(k!(40-k)!)]\,(1/4)^k (3/4)^{40-k}.$$

Ci interessa la probabilità di ottenere almeno cinque centri, che si esprime con la formula

$$P = \sum_{k=5}^{40} P_k .$$

È però più semplice calcolarla mediante $P = 1 - Q$, dove

$$Q = \sum_{k=0}^{4} P_k$$

è la probabilità di ottenere meno di cinque centri.

Si può calcolare che

$$P_0 = \left(\frac{3}{4}\right)^{40} \sim 0.00001,$$

$$P_1 = 40 \left(\frac{3}{4}\right)^{39} \frac{1}{4} \sim 0.00013,$$

$$P_2 = \frac{40 \cdot 39}{2} \left(\frac{3}{4}\right)^{38} \left(\frac{1}{4}\right)^2 \sim 0.00087,$$

$$P_3 = \frac{40 \cdot 39 \cdot 38}{2 \cdot 3} \left(\frac{3}{4}\right)^{37} \left(\frac{1}{4}\right)^3 \sim 0.0037,$$

$$P_4 = \frac{40 \cdot 39 \cdot 38 \cdot 37}{2 \cdot 3 \cdot 4} \left(\frac{3}{4}\right)^{36} \left(\frac{1}{4}\right)^4 \sim 0.0113,$$

cosicché

$$Q \approx 0{,}016,\ P \approx 0{,}984.$$

La probabilità P così ottenuta è un po' più vicina alla certezza di quanto abitualmente si ritenga sufficiente nella teoria del tiro. Di solito si considera ammissibile determinare il numero di proiettili necessario ad assicurare il risultato con probabilità 0,95.

L'esempio precedente è schematico, ma mostra con sufficiente dettaglio l'importanza pratica del calcolo delle probabilità. Dopo aver stabilito sperimentalmente la dipendenza degli scarti probabili dalla distanza, senza dover sparare un numero enorme di colpi, abbiamo potuto rispondere mediante semplici calcoli a domande molto diverse. Lo stesso avviene negli altri campi in cui l'influenza congiunta di molti fattori casuali conduce a una legge statistica. L'esame diretto della massa delle osservazioni mette in luce soltanto le leggi statistiche più semplici e alcune probabilità fondamentali. Mediante le leggi della teoria della probabilità possiamo poi usare queste probabilità elementari per calcolare quelle di eventi più complessi e dedurne le leggi statistiche.

Talvolta si riesce a evitare del tutto la raccolta di grandi masse di dati, perché le probabilità possono essere determinate da considerazioni di simmetria sufficientemente convincenti. La conclusione tradizionale che un dado, ossia un cubo di materiale omogeneo, lanciato a un'altezza sufficiente, cada su ciascuna faccia con la stessa probabilità fu raggiunta molto prima che si raccogliessero sistematicamente dati per verificarla. Esperimenti sistematici di questo tipo furono compiuti negli ultimi tre secoli, soprattutto dagli autori di manuali, quando la teoria della probabilità era già una scienza sviluppata. I risultati furono soddisfacenti, ma la loro estensione a casi analoghi suscita scarso interesse. Per esempio, per quanto ci risulta, nessuno ha compiuto esperimenti abbastanza estesi lanciando dadi omogenei regolari a dodici facce. Non vi è però motivo di dubitare che, effettuando 12 000 lanci, ciascuna faccia comparirebbe circa mille volte.

Le probabilità fondamentali ricavate da argomenti di simmetria o omogeneità svolgono un ruolo importante anche in problemi scientifici di grande rilievo: per esempio nelle collisioni o negli incontri ravvicinati fra molecole in moto casuale in un gas, oppure nel moto delle stelle in una galassia, dove si sono ottenuti successi altrettanto notevoli. Naturalmente, in questi casi più delicati si preferisce controllare le ipotesi teoriche confrontandole con osservazioni o esperimenti.

§3. La legge dei grandi numeri e i teoremi limite

È del tutto naturale desiderare una maggiore precisione quantitativa nell'affermazione che, in una serie «lunga» di prove, la frequenza di un evento si «avvicina» alla sua probabilità. Dobbiamo però comprendere la delicatezza del problema. Nei casi più tipici della teoria della probabilità, anche in una serie arbitrariamente lunga di prove resta teoricamente possibile ottenere uno dei due valori estremi della frequenza:

$$\mu/n = n/n = 1 \text{ e } \mu/n = 0/n = 0.$$

Qualunque sia il numero n delle prove, non si può quindi affermare con assoluta certezza che valga, per esempio, la disuguaglianza

$$\left|\frac{\mu}{n} - p\right| < \frac{1}{10}.$$

Se A è, per esempio, l'uscita del sei nel lancio di un dado, in n prove indipendenti la probabilità di ottenere sei in tutte le prove è $(\frac{1}{6})^n > 0$. In altre parole, con probabilità $(\frac{1}{6})^n$ la frequenza del sei sarà uguale a uno; con probabilità $(1 - \frac{1}{6})^n > 0$ il sei non uscirà mai, e la sua frequenza sarà zero.

In tutti i problemi analoghi, una stima non banale della vicinanza fra frequenza e probabilità non può valere con assoluta certezza, ma soltanto con una probabilità minore di uno. Si può mostrare, per esempio, che in prove indipendenti,[5] con probabilità costante p dell'evento in ogni prova, la disuguaglianza

$$\left|\frac{\mu}{n} - p\right| < 0.02 \qquad (7)$$

per la frequenza μ/n, con n = 10 000 e per ogni p, è soddisfatta con probabilità

$$P > 0.9999. \qquad (8)$$

Vogliamo anzitutto sottolineare che, in questa formulazione, la stima quantitativa della vicinanza fra μ/n e p richiede l'introduzione di una nuova probabilità P.

Il significato pratico della stima (8) è questo: se compiamo N serie, ciascuna di n prove, e contiamo le M serie in cui vale la (7), per N sufficientemente grande avremo approssimativamente

$$\frac{M}{N} \approx P > 0.9999. \qquad (9)$$

Ma, se vogliamo precisare la (9), sia riguardo alla vicinanza di M/N a P, sia riguardo alla fiducia con cui possiamo affermarne la validità, dobbiamo ricorrere a considerazioni generali dello stesso tipo di quelle introdotte per chiarire la vicinanza fra μ/n e p. Possiamo ripeterle

quante volte vogliamo; è però chiaro che questo procedimento non ci libererà mai dalla necessità, all'ultimo passaggio, di riferirci alle probabilità nel senso originario, non precisato, del termine.

Sarebbe sbagliato pensare che difficoltà di questo tipo siano peculiari della teoria della probabilità. Nello studio matematico di eventi reali costruiamo sempre un modello. Le discrepanze fra il corso effettivo degli eventi e il modello teorico possono a loro volta diventare oggetto di studio matematico. Ma anche per esse dobbiamo costruire un modello, che useremo senza analizzare formalmente le ulteriori discrepanze che emergerebbero nell'esperimento reale.

Osserviamo inoltre che, applicando concretamente la stima[6]

$$\mathbf{P}\left\{\left|\frac{\mu}{n} - p\right| < 0.02\right\} > 0.9999 \qquad (10)$$

a una singola serie di n prove, facciamo già affidamento su considerazioni di simmetria. La (10) dice che, per un numero molto grande N di serie, la (7) sarà soddisfatta in almeno il 99,99% dei casi. È allora naturale attendersi con grande fiducia che essa valga anche per la particolare serie di n prove che ci interessa. Possiamo però attendercelo solo se abbiamo ragione di supporre che tale serie non occupi una posizione speciale fra le altre, ossia non presenti caratteristiche peculiari.

Le probabilità che possiamo decidere di trascurare variano secondo la situazione pratica. Abbiamo osservato che, nel calcolo preliminare del numero di proiettili necessario a ottenere un risultato, si assume come criterio una probabilità di successo di 0,95: le probabilità trascurate non superano dunque 0,05. La ragione è che, se trascurassimo soltanto una probabilità di 0,01, per esempio, richiederemmo necessariamente molti più proiettili; in pratica potremmo allora concludere che il compito non è realizzabile nel tempo disponibile o con le munizioni assegnate.

Anche nelle ricerche scientifiche siamo talvolta costretti a usare metodi statistici che trascurano probabilità dell'ordine di 0,05, sebbene ciò, secondo il criterio qui adottato, debba avvenire soltanto quando raccogliere dati più estesi è molto difficile. Consideriamo il seguente esempio. Si supponga che, in condizioni specifiche, il farmaco abitualmente usato contro una certa malattia dia risultati positivi nel 50% dei casi, cioè con probabilità 0,5. Viene proposto un nuovo preparato e, per verificarne il vantaggio, si decide di usarlo in dieci casi scelti senza distorsioni fra i malati. Si conviene di dichiarare dimostrato il vantaggio se il risultato è positivo in almeno otto casi su dieci. È facile calcolare che questa procedura comporta una probabilità dell'ordine di 0,05 di concludere erroneamente a favore del nuovo preparato quando esso è altrettanto efficace del precedente, o persino meno efficace. Se le dieci prove sono indipendenti e la probabilità di esito positivo in ciascuna è p, le probabilità di ottenere rispettivamente 10, 9 e 8 esiti positivi sono

$$P_{10} = p^{10}, \quad P_9 = 10p^9(1-p), \quad P_8 = 45p^8(1-p)^2.$$

Nel caso $p = \frac{1}{2}$, la loro somma è

$$P = P_{10} + P_9 + P_8 = \frac{56}{1024} \sim 0.05.$$

Dunque, supponendo che il nuovo preparato abbia esattamente la stessa efficacia del precedente, corriamo un rischio dell'ordine di 0,05 di dichiararlo erroneamente migliore. Per ridurre tale probabilità a circa 0,01 senza aumentare il numero $n = 10$ delle prove, occorre richiedere almeno nove esiti positivi su dieci. Se questo requisito appare troppo severo ai sostenitori del preparato, bisogna aumentare sensibilmente il numero delle prove. Per esempio, con $n = 100$ e richiedendo più di 65 risultati positivi, cioè $\mu > 65$, la probabilità di questo errore sotto l'ipotesi $p = 0{,}5$ è circa $P \approx 0{,}000895$.

Per ricerche scientifiche rigorose, sostiene il testo, il criterio di 0,05 è insufficiente; tuttavia anche in studi accurati, come l'elaborazione delle osservazioni astronomiche, si trascurano abitualmente probabilità d'errore di 0,001 o 0,003. D'altra parte alcuni risultati scientifici fondati sulle leggi probabilistiche sono assai più affidabili, nel senso che comportano il trascurare probabilità ancora minori. Torneremo sulla questione.

Negli esempi precedenti abbiamo usato casi particolari della formula binomiale (6)

$$P_m = C_n^m p^m (1-p)^{n-m}$$

per la probabilità di ottenere esattamente m risultati positivi in n prove indipendenti, ciascuna con probabilità p di successo. Consideriamo, mediante questa formula, la questione posta all'inizio del paragrafo: determinare la probabilità

$$P = \mathbf{P}\left\{\left|\frac{\mu}{n} - p\right| < \epsilon\right\}, \qquad (11)$$

dove μ è il numero effettivo degli esiti positivi.[7] Essa si può evidentemente scrivere come somma delle P_m per cui m soddisfa la disuguaglianza

$$\left|\frac{m}{n} - p\right| < \epsilon, \qquad (12)$$

cioè nella forma

$$P = \sum_{m=m_1}^{m_2} P_m, \qquad (13)$$

dove m_1 e m_2 sono rispettivamente il più piccolo e il più grande valore di m che soddisfano la (12).

Per n piuttosto grande, la (13) è poco comoda per il calcolo diretto. Ciò spiega l'importanza della formula asintotica scoperta da de Moivre per $p = \frac{1}{2}$ e da Laplace per p generico. Essa permette di trovare assai semplicemente P_m e di studiarne il comportamento per grandi n:

$$P_m \sim \frac{1}{\sqrt{2\pi np(1-p)}}\, e^{-(m-np)^2/2np(1-p)}. \qquad (14)$$

Se p non è troppo vicino a zero o a uno, nella regione centrale la formula è già sufficientemente precisa per n dell'ordine di 100. Ponendo

$$t = \frac{m - np}{\sqrt{np(1-p)}}, \qquad (15)$$

la (14) diventa

$$P_m \sim \frac{1}{\sqrt{2\pi np(1-p)}}\, e^{-t^2/2}. \qquad (16)$$

Dalle (13) e (16) si può ricavare una rappresentazione approssimata della probabilità (11):

$$P \sim \frac{1}{\sqrt{2\pi}} \int_{-T}^{T} e^{-t^2/2}\, dt = F(T), \qquad (17)$$

dove

$$T = \epsilon \sqrt{\frac{n}{p(1-p)}}. \qquad (18)$$

Per p fissato e diverso da zero e da uno, la differenza fra i due membri della (17) tende a zero, uniformemente rispetto a ϵ, quando $n \to \infty$. Della funzione $F(T)$ sono state costruite tavole dettagliate. Eccone un piccolo estratto:

T	1	2	3	4
F	0.68269	0.95450	0.99730	0.99993

Per $T \to \infty$, i valori di $F(T)$ tendono a uno.

Dalla (17) ricaviamo una stima della probabilità

$$\mathbf{P}\left\{\left|\frac{\mu}{n} - p\right| < 0.02\right\}$$

per n = 10 000. Poiché

$$T = \frac{2}{\sqrt{p(1-p)}},$$

abbiamo

$$P \approx F\left(\frac{2}{\sqrt{p(1-p)}}\right).$$

La funzione $F(T)$ è crescente; per ottenere una stima inferiore di P indipendente da p, dobbiamo quindi prendere il più piccolo valore possibile di T al variare di p. Esso si ottiene per $p = \frac{1}{2}$ ed è uguale a 4. Dunque, nell'approssimazione normale,

$$P \approx F(T),\ F(T) \geq F(4) \approx 0{,}99994. \quad (19)$$

Nella (19) non si tiene conto dell'errore dovuto al carattere approssimato della (17). Stimando tale errore si può mostrare che, in ogni caso, $P > 0{,}9999$.

A proposito di questa applicazione, va osservato che le stime del resto nella (17), fornite dalle opere teoriche sulla probabilità, furono a lungo insoddisfacenti. Le applicazioni della (17) e di formule analoghe a piccoli valori di n, oppure a probabilità p molto vicine a zero o a uno, spesso particolarmente importanti, si fondavano quindi soltanto su verifiche sperimentali in un numero limitato di esempi, senza valide stime dell'errore possibile. Ricerche più approfondite mostrarono inoltre che, in molti casi pratici importanti, occorrono sia stime del resto sia ulteriori perfezionamenti delle formule asintotiche; altrimenti il resto sarebbe troppo grande. In entrambe le direzioni i risultati più completi sono dovuti a S. N. Bernštejn.

Le relazioni (11), (17) e (18) si possono riscrivere nella forma

$$\mathbf{P}\left\{\left|\frac{\mu}{n} - p\right| < t\sqrt{\frac{p(1-p)}{n}}\right\} \sim F(t). \quad (20)$$

Per t sufficientemente grande, il secondo membro della (20), che non contiene n, è arbitrariamente vicino a uno, ossia alla probabilità della certezza. Vediamo così che, di regola, lo scarto della frequenza μ/n dalla probabilità p è dell'ordine di $1/\sqrt{n}$. Questa relazione fra precisione di una legge probabilistica e radice quadrata del numero delle osservazioni è tipica di molti altri problemi. Nelle esposizioni divulgative si afferma talvolta che la «legge della radice quadrata di n» è la legge fondamentale della probabilità. L'idea fu resa precisa dal grande matematico russo P. L. Čebyšev, con l'introduzione e l'uso sistematico dei concetti di «speranza matematica» e «varianza» per somme e medie aritmetiche di «variabili aleatorie».

Si chiama variabile aleatoria una quantità che, sotto condizioni assegnate S, può assumere valori diversi con probabilità determinate. Per i nostri scopi basta considerare variabili con un numero finito di valori. Per assegnare la cosiddetta distribuzione di probabilità di una tale variabile ξ, basta indicarne i valori possibili $x_1, x_2, \ldots, x_s$ e le probabilità

$$P_r = P\{\xi = x_r\},\ r = 1, \ldots, s.$$

La somma di queste probabilità, estesa a tutti i valori possibili di ξ, è sempre uguale a uno:

$$\sum_{r=1}^{s} P_r = 1.$$

Il numero degli esiti positivi in n esperimenti, studiato sopra, è un esempio di variabile aleatoria.

La speranza matematica, o valore atteso, della variabile ξ è l'espressione

$$\mathbf{M}(\xi) = \sum_{r=1}^{s} P_r x_r ,$$

La varianza di ξ è la speranza matematica del quadrato dello scarto $\xi - \mathrm{M}(\xi)$, cioè

$$\mathbf{D}(\xi) = \sum_{r=1}^{s} P_r [x_r - \mathbf{M}(\xi)]^2 .$$

La radice quadrata della varianza

$$\sigma_\xi = \sqrt{\mathrm{D}(\xi)}$$

si chiama deviazione standard, o scarto quadratico medio, della variabile rispetto alla sua speranza matematica $\mathrm{M}(\xi)$.

Alla base delle applicazioni più semplici della varianza e della deviazione standard vi è la celebre disuguaglianza di Čebyšev, per $t > 0$:

$$\mathbf{P}\{|\xi - \mathbf{M}(\xi)| \leqslant t\,\sigma_\xi\} \geqslant 1 - \frac{1}{t^2} . \qquad (21)$$

Essa mostra che gli scarti di ξ da $\mathrm{M}(\xi)$ molto maggiori di σ_ξ sono rari.

Per una somma di variabili aleatorie

$$\xi = \xi^{(1)} + \xi^{(2)} + \cdots + \xi^{(n)},$$

le speranze matematiche soddisfano sempre l'uguaglianza

$$\mathbf{M}(\xi) = \mathbf{M}(\xi^{(1)}) + \mathbf{M}(\xi^{(2)}) + \cdots + \mathbf{M}(\xi^{(n)}). \qquad (22)$$

L'uguaglianza analoga per la varianza,

$$\mathbf{D}(\xi) = \mathbf{D}(\xi^{(1)}) + \mathbf{D}(\xi^{(2)}) + \cdots + \mathbf{D}(\xi^{(n)}) \qquad (23)$$

vale invece soltanto sotto determinate condizioni. È sufficiente, per esempio, che le variabili $\xi^{(i)}$ e $\xi^{(j)}$ con indici diversi non siano «correlate», cioè che per $i \neq j$ valga l'uguaglianza[8]

$$\mathbf{M}\{[\xi^{(i)} - \mathbf{M}(\xi^{(i)})]\,[\xi^{(j)} - \mathbf{M}(\xi^{(j)})]\} = 0 \qquad (24)$$

appena scritta.

In particolare la (24) vale se $\xi^{(i)}$ e $\xi^{(j)}$ sono indipendenti.[9] Per addendi mutuamente indipendenti, dunque, la (23) vale sempre. Per la media aritmetica

$$\zeta = \frac{1}{n}(\xi^{(1)} + \xi^{(2)} + \cdots + \xi^{(n)})$$

dalla (23) segue

$$\mathbf{D}(\zeta) = \frac{1}{n^2}\left[\mathbf{D}(\xi^{(1)}) + \mathbf{D}(\xi^{(2)}) + \cdots + \mathbf{D}(\xi^{(n)})\right]. \qquad (25)$$

Supponiamo ora che la varianza di ciascun addendo non superi una stessa costante:

$$\mathbf{D}(\xi^{(i)}) \leqslant C^2.$$

Allora dalla (25)

$$\mathbf{D}(\zeta) \leqslant \frac{C^2}{n},$$

e dalla disuguaglianza di Čebyšev, per ogni $t > 0$,

$$\mathbf{P}\left\{|\zeta - \mathbf{M}(\zeta)| \leqslant \frac{tC}{\sqrt{n}}\right\} \geqslant 1 - \frac{1}{t^2}. \qquad (26)$$

La (26) esprime la cosiddetta legge dei grandi numeri nella forma stabilita da Čebyšev: se le variabili $\xi^{(i)}$ sono mutuamente indipendenti e le loro varianze sono limitate da una stessa costante, al crescere di n la media aritmetica ζ si discosta sempre più raramente dalla speranza matematica $\mathrm{M}(\zeta)$.

Più precisamente, si dice che la successione di variabili

$$\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(n)}, \ldots$$

soddisfa la legge dei grandi numeri se, per le corrispondenti medie aritmetiche ζ e per ogni costante $\epsilon > 0$,

$$\mathbf{P}\{|\zeta - \mathbf{M}(\zeta)| \leqslant \epsilon\} \to 1 \qquad (27)$$

quando $n \to \infty$.

Per passare dalla disuguaglianza (26) alla relazione limite (27) basta porre, con $C > 0$,

$$t = \epsilon\frac{\sqrt{n}}{C}.$$

Molte ricerche di A. A. Markov, S. N. Bernštejn, A. Ja. Hinčin e altri furono dedicate ad ampliare il più possibile le condizioni di validità della (27), cioè della legge dei grandi numeri. Esse hanno un'importanza teorica fondamentale; ancora più importante è però lo studio preciso della distribuzione di probabilità dello scarto della media, $\zeta - \mathrm{M}(\zeta)$.

Uno dei maggiori contributi della scuola matematica russa classica alla teoria della probabilità fu stabilire che, sotto condizioni molto ampie, la relazione

$$\mathbf{P}\{t_1\sigma_\zeta < \zeta - \mathbf{M}(\zeta) < t_2\sigma_\zeta\} \sim \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-t^2/2}\, dt \qquad (28)$$

vale asintoticamente, cioè con precisione crescente quando n aumenta indefinitamente.

Čebyšev ne diede una dimostrazione quasi completa per addendi indipendenti e limitati, sotto le condizioni di non degenerazione richieste dal teorema. Markov colmò una lacuna nel ragionamento e ampliò le condizioni di applicabilità della (28). Condizioni ancora più generali furono date da Ljapunov. L'applicabilità della (28) a somme di addendi dipendenti fu studiata in modo particolarmente completo da S. N. Bernštejn.

La (28) comprende un numero tanto grande di casi particolari che da tempo viene chiamata teorema centrale del limite. Sebbene ricerche più recenti l'abbiano inclusa in una serie di leggi più generali, la sua importanza resta difficilmente sopravvalutabile.

Se gli addendi sono indipendenti e hanno tutti la stessa varianza,

$$\mathbf{D}(\xi^{(i)}) = \sigma^2,$$

sotto le ipotesi del teorema centrale del limite, è comodo usare la (25) per scrivere la (28) nella forma

$$\mathbf{P}\left\{\frac{t_1\sigma}{\sqrt{n}} < \zeta - \mathbf{M}(\zeta) < \frac{t_2\sigma}{\sqrt{n}}\right\} \sim \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-t^2/2}\, dt. \qquad (29)$$

Mostriamo che la (29) contiene la soluzione del problema, già considerato, di valutare lo scarto della frequenza μ/n dalla probabilità p. Introduciamo a questo scopo le variabili aleatorie $\xi^{(i)}$ così definite:

$\xi^{(i)} = 0$, se l'i-esima prova ha esito negativo;
$\xi^{(i)} = 1$, se l'i-esima prova ha esito positivo.

Si verifica facilmente che

$$\mu = \xi^{(1)} + \xi^{(2)} + \cdots + \xi^{(n)}, \qquad \frac{\mu}{n} = \zeta,$$

$$\mathbf{M}(\xi^{(i)}) = p, \quad \mathbf{D}(\xi^{(1)}) = p(1-p), \quad \mathbf{M}(\zeta) = p,$$

e la (29) dà

$$\mathbf{P}\left\{t_1\sqrt{\frac{p(1-p)}{n}} < \frac{\mu}{n} - p < t_2\sqrt{\frac{p(1-p)}{n}}\right\} \sim \frac{1}{\sqrt{2\pi}} \int_{t_1}^{t_2} e^{-t^2/2}\, dt,$$

Ponendo $t_1 = -t$ e $t_2 = t$, si ritrova la (20).

§4. Ulteriori osservazioni sui concetti fondamentali della probabilità

Nel §1, parlando di eventi casuali le cui frequenze tendono a stabilizzarsi, cioè a raggrupparsi, in lunghe successioni di esperimenti ripetuti in condizioni costanti, intorno a un valore tipico chiamato probabilità P(*A*/*S*), siamo stati imprecisi sotto due aspetti. Anzitutto non abbiamo indicato quanto debba essere lunga una successione di n_r esperimenti per rendere indubbia la presunta stabilità. In altre parole, non abbiamo detto quali scarti delle frequenze μ_r/n_r, fra loro o dal valore tipico *p*, siano ammissibili per serie di lunghezze assegnate n_1, n_2, ..., n_s. Questa imprecisione è inevitabile nella prima formulazione dei concetti di una scienza nuova. Non è maggiore della ben nota indeterminatezza che circonda i concetti geometrici più semplici di punto e retta e il loro significato fisico. Questo aspetto è stato chiarito nel §3.

Più fondamentale è la seconda mancanza di chiarezza: essa riguarda il modo di formare le successioni di prove nelle quali si deve esaminare la stabilità della frequenza di *A*.

Come si è detto, ricorriamo ai metodi statistici e probabilistici quando una previsione precisa del corso degli eventi è impossibile. Ma, se vogliamo costruire artificialmente una successione di eventi quanto più possibile puramente casuale, dobbiamo assicurarci che non vi siano metodi per individuare in anticipo i casi in cui *A* tenderà a verificarsi con frequenza superiore al normale.

Precauzioni simili si adottano nell'organizzazione delle lotterie statali. Se un'estrazione assegna *M* biglietti vincenti fra *N* biglietti, la probabilità di vincere per un singolo biglietto è $p = M/N$. Ciò significa che, comunque si scelga prima dell'estrazione un insieme sufficientemente numeroso di *n* biglietti, si può essere praticamente certi che il rapporto μ/n fra i biglietti vincenti μ e il totale *n* del gruppo sarà vicino a *p*. Chi preferisce numeri pari non avrà, per esempio, un vantaggio sistematico su chi preferisce numeri dispari. Analogamente non vi sarà vantaggio nello scegliere numeri con esattamente tre fattori primi, oppure numeri vicini a quelli vincenti nella lotteria precedente, e così via.

Analogamente, usando un cannone ben costruito di un certo tipo, un equipaggio addestrato e proiettili sottoposti al consueto controllo di qualità, lo scarto dei punti d'impatto dalla posizione media sarà, in valore assoluto, minore dello scarto probabile *B* in circa metà dei casi. La stessa frazione si osserva sia in una serie di prove consecutive, sia contando separatamente i colpi di numero pari e quelli di numero dispari nell'ordine di tiro. Se però scegliessimo proiettili particolarmente omogenei per peso e altre caratteristiche, la dispersione potrebbe diminuire sensibilmente: avremmo allora una serie in cui la frazione degli scarti maggiori del valore standard *B* sarebbe molto inferiore alla metà.

Perciò si può dire che un evento *A* è «casuale» o «stocastico» e assegnargli una probabilità determinata

$$p = \mathbf{P}(A/S)$$

soltanto dopo aver stabilito la classe dei modi ammissibili di costruire le serie di esperimenti. Si supporrà che la natura di tale classe sia inclusa nelle condizioni *S*.

Per condizioni *S* assegnate, la casualità di *A* e il suo possedere la probabilità $p = \mathrm{P}(A/S)$ esprimono il carattere oggettivo del legame fra *S* e *A*. Non esiste, in altre parole, un evento assolutamente casuale: un evento è casuale o predeterminato secondo la relazione in cui lo si considera. In condizioni specifiche, però, può essere casuale in un senso del tutto non soggettivo, indipendente dalle conoscenze di qualsiasi osservatore. Immaginiamo un osservatore capace di conoscere ogni proprietà e circostanza particolare del volo dei proiettili, e quindi di prevedere per ciascuno lo scarto dalla traiettoria media. La sua presenza non impedirebbe ai proiettili di disperdersi secondo le leggi della probabilità, purché il tiro avvenisse nel modo consueto e non secondo le istruzioni dell'osservatore immaginario.

Va osservato che nel mondo reale le serie in cui le frequenze tendono a stabilizzarsi intorno al valore normale, cioè alla probabilità, si formano del tutto indipendentemente dal nostro intervento. Proprio per il carattere casuale del moto delle molecole di un gas, per esempio, il numero di molecole che in un intervallo di tempo anche molto breve colpiscono una piccola porzione prefissata della parete del recipiente, o della superficie di un corpo immerso nel gas, risulta con grande precisione proporzionale all'area della porzione e alla durata dell'intervallo. Quando il numero di urti è piccolo, anche gli scarti da questa proporzionalità seguono le leggi della probabilità e producono fenomeni come il moto browniano, di cui parleremo più avanti.

Passiamo ora al significato oggettivo dell'indipendenza. Ricordiamo che la probabilità condizionata di *A* dato *B*, con $\mathrm{P}(B) > 0$, è definita dalla formula

$$\mathbf{P}(A/B) = \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}. \qquad (30)$$

Ricordiamo anche che *A* e *B* si dicono indipendenti se, come nella (4),

$$\mathbf{P}(AB) = \mathbf{P}(A)\,\mathbf{P}(B).$$

Dalla loro indipendenza e da $\mathrm{P}(B) > 0$ segue

$$\mathbf{P}(A/B) = \mathbf{P}(A).$$

Tutti i teoremi della teoria matematica della probabilità relativi a eventi indipendenti valgono per qualsiasi evento che soddisfi la (4), o la sua generalizzazione all'indipendenza congiunta di più eventi. Questi teoremi avrebbero tuttavia scarso interesse se tale definizione non fosse collegata alle proprietà di eventi oggettivi indipendenti in senso causale.

È noto, per esempio, che la probabilità di una nascita maschile presenta una sufficiente stabilità; nel dato qui assunto è $\mathrm{P}(A) = 22/43$. Se *B* indica la condizione che la nascita avvenga in un giorno di congiunzione fra Giove e Marte, supponendo che la posizione dei pianeti non influisca sul destino degli individui, la probabilità condizionata ha lo stesso valore: $\mathrm{P}(A/B) =$

22/43. Un conteggio effettivo delle frequenze delle nascite maschili in tali particolari condizioni astrologiche darebbe quindi approssimativamente la medesima frequenza. Benché un conteggio simile probabilmente non sia mai stato effettuato su scala abbastanza ampia, non vi è motivo di dubitare del risultato.

Questo esempio, tratto da un argomento ormai superato, serve a mostrare che lo sviluppo delle conoscenze umane consiste sia nello stabilire relazioni valide fra i fenomeni, sia nel confutare relazioni immaginarie: ossia, quando è opportuno, nello stabilire l'indipendenza di due insiemi di eventi. Smascherare i tentativi privi di fondamento degli astrologi di collegare eventi che in realtà non sono connessi è uno degli esempi classici.

Naturalmente non bisogna trattare il concetto di indipendenza in modo troppo assoluto. Per la legge di gravitazione universale, per esempio, è indubbio che i moti dei satelliti di Giove esercitino una certa influenza anche sul volo di un proiettile. È però altrettanto evidente che in pratica tale influenza si può trascurare. Dal punto di vista filosofico, in una situazione concreta sarebbe forse più appropriato parlare di irrilevanza della dipendenza anziché di indipendenza. Comunque sia, l'indipendenza degli eventi, nel senso concreto e relativo appena precisato, non contraddice il principio dell'interconnessione universale dei fenomeni: ne costituisce un necessario complemento.

Calcolare probabilità mediante formule ricavate dall'indipendenza conserva interesse pratico anche quando eventi inizialmente indipendenti diventano dipendenti in seguito al loro stesso svolgersi. Si possono, per esempio, calcolare probabilità di collisione fra particelle della radiazione cosmica e particelle del mezzo attraversato, supponendo che il moto delle particelle del mezzo, fino all'arrivo nelle loro vicinanze di una particella cosmica veloce, proceda indipendentemente da quest'ultima. Si può calcolare la probabilità che un proiettile nemico colpisca la pala di un'elica in rotazione supponendo che la posizione della pala rispetto all'asse non dipenda dalla traiettoria del proiettile. L'ipotesi sarebbe naturalmente falsa per i proiettili sparati dall'aviatore stesso, sincronizzati per passare fra le pale. Esempi simili si possono moltiplicare senza limite.

Si può persino affermare che, ovunque le leggi probabilistiche si manifestino chiaramente, interviene un gran numero di fattori che, se non sono del tutto indipendenti, sono collegati fra loro soltanto debolmente.

Ciò non significa affatto che si debbano introdurre senza spirito critico ipotesi di indipendenza. Al contrario, occorre anzitutto scegliere con particolare cura i criteri con cui verificarle; inoltre bisogna studiare attentamente i casi limite in cui una dipendenza va ammessa, ma complica le leggi probabilistiche pertinenti. Abbiamo già ricordato le ampie ricerche condotte in questa direzione dalla scuola russa classica.

Concludiamo la discussione osservando che, come la definizione (4) per due eventi, anche la definizione formale di indipendenza di più variabili aleatorie è assai più ampia del concetto pratico di indipendenza come assenza di legame causale.

Supponiamo, per esempio, che un punto ξ cada nell'intervallo $[0, 1]$ in modo tale che, per

$$0 \leqslant a \leqslant b \leqslant 1$$

la probabilità di appartenere al segmento $[a, b]$ sia uguale alla sua lunghezza $b - a$. È facile dimostrare che, nello sviluppo decimale

$$\xi = \frac{\alpha_1}{10} + \frac{\alpha_2}{100} + \frac{\alpha_3}{1000} + \cdots$$

dell'ascissa di ξ, le cifre α_k sono mutuamente indipendenti, pur essendo collegate dal modo in cui vengono prodotte.[10] Da questo fatto seguono molti risultati teorici, alcuni dei quali hanno interesse pratico.

Una simile flessibilità della definizione formale non è un difetto: amplia il campo di applicazione dei teoremi stabiliti sotto ipotesi di indipendenza. Essi valgono sia quando l'indipendenza è postulata per considerazioni pratiche, sia quando è dimostrata mediante calcoli a partire da precedenti ipotesi sulle distribuzioni di probabilità degli eventi e delle variabili studiate.

In generale, lo studio della struttura formale dell'apparato matematico della probabilità ha condotto a risultati interessanti. Si è visto che questo apparato occupa un posto ben determinato nella classificazione degli oggetti fondamentali della matematica contemporanea, i cui contorni vanno chiarendosi gradualmente.

Abbiamo già parlato dell'intersezione AB e dell'unione $A \cup B$ di due eventi. Ricordiamo che essi sono incompatibili se la loro intersezione è vuota, cioè se $AB = N$, dove N indica l'evento impossibile.

L'assioma fondamentale della teoria elementare richiede, come nel §2, che, se $AB = N$, valga

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B).$$

I concetti fondamentali della probabilità, ossia gli eventi casuali e le loro probabilità, hanno proprietà del tutto analoghe a quelle delle figure piane e delle loro aree. Basta intendere per AB l'intersezione, o parte comune, di due figure; per $A \cup B$ la loro unione; per N la figura convenzionalmente «vuota»; e per $\mathrm{P}(A)$ l'area di A. L'analogia è allora completa.

Le stesse osservazioni valgono per i volumi delle figure tridimensionali.

La teoria più generale di enti di questo tipo, che comprende come casi particolari le teorie dell'area e del volume, è oggi comunemente chiamata teoria della misura. Ne parleremo nel capitolo XV, in relazione alla teoria delle funzioni di variabile reale.

Rispetto alla teoria generale della misura, e in particolare a quelle di area e volume, la probabilità presenta una caratteristica speciale: non supera mai uno. Tale valore massimo spetta all'evento certo U:

$$\mathbf{P}(U) = 1.$$

L'analogia non è affatto superficiale. L'intera teoria matematica della probabilità si può costruire formalmente come teoria della misura, con la speciale ipotesi che la misura dell'«intero spazio» U sia uno.[11]

Questo punto di vista ha chiarito completamente la costruzione formale della teoria della probabilità e ha prodotto progressi concreti sia in essa sia in teorie formalmente affini. In probabilità si sono ottenuti successi grazie ai raffinati metodi della teoria metrica delle funzioni di variabile reale; al tempo stesso, i metodi probabilistici si sono rivelati applicabili a campi vicini della matematica mediante un trasferimento formale e rigoroso, non per semplice analogia. Ovunque siano soddisfatti gli assiomi della probabilità, se ne possono applicare le conseguenze, anche se il campo considerato non ha nulla a che fare con la casualità del mondo reale.

L'esistenza di una teoria assiomatica ci preserva dalla tentazione di «definire» la probabilità con metodi che pretendano di costruire una teoria rigorosa e puramente formale partendo da caratteristiche immediatamente suggerite dalle scienze naturali. Tali definizioni corrispondono grosso modo alla «definizione» geometrica di punto come risultato di un'infinita serie di riduzioni di un corpo fisico, ciascuna delle quali ne dimezza il diametro.

In definizioni di questo genere, la probabilità viene assunta come limite della frequenza quando il numero degli esperimenti cresce indefinitamente. Ma l'ipotesi stessa del carattere probabilistico degli esperimenti, cioè della tendenza delle frequenze a raggrupparsi intorno a un valore costante, rimarrebbe valida, come la «casualità» di ciascun evento, soltanto se determinate condizioni si conservassero con precisione assoluta per un tempo illimitato. Il passaggio esatto al limite

$$\frac{\mu}{n} \to p$$

non può dunque avere un significato oggettivo diretto. Formulare la stabilità delle frequenze mediante questo limite richiede di definire i modi ammissibili di costruire una successione infinita di esperimenti, cosa possibile soltanto attraverso una finzione matematica. Questo insieme di concetti meriterebbe seria considerazione se il risultato fosse una teoria tanto peculiare da non poter essere fondata rigorosamente in altro modo. Ma, come si è detto, basta fondare la teoria della probabilità sulla moderna teoria della misura, aggiungendo la condizione

$$\mathbf{P}(U) = 1.$$

In generale, per un'analisi pratica del concetto di probabilità non occorre riferirsi alla sua definizione formale. Sul suo significato applicativo possiamo dire questo: $P(A/S)$ è un numero intorno al quale le frequenze tendono a raggrupparsi sotto condizioni S che determinano il modo ammissibile di organizzare gli esperimenti. Tale raggruppamento si fa più preciso all'aumentare degli esperimenti, sempre condotti conservando le condizioni iniziali, e raggiunge un grado soddisfacente di affidabilità e precisione entro un numero di prove concretamente realizzabile.

Il problema importante nella pratica non è dare una definizione formalmente precisa della casualità, ma chiarire il più ampiamente possibile le condizioni nelle quali compare la casualità del tipo descritto. Bisogna comprendere che le ipotesi sul carattere probabilistico di un fenomeno si fondano assai raramente su una verifica statistica diretta. Soltanto nella fase iniziale dell'introduzione dei metodi probabilistici in un nuovo campo si procede mediante l'osservazione puramente empirica della costanza delle frequenze. Dal §3 vediamo che, per una precisione ε, occorre un numero di prove dell'ordine di $n = 1/\varepsilon^2$, a parità degli altri criteri. Per esempio, per stabilire una probabilità con precisione dell'ordine di 0,0001 occorrono serie dell'ordine di 100 000 000 di prove ciascuna.

L'ipotesi di casualità probabilistica viene introdotta molto più spesso in base a considerazioni di simmetria o di successione degli eventi, e poi verificata indirettamente. Per esempio, poiché il numero delle molecole in un volume macroscopico di gas è dell'ordine di 10^{20} o più, il numero $\sqrt{n}$ che interviene nelle deduzioni probabilistiche della teoria cinetica è molto grande; molte di queste deduzioni sono quindi verificate con grande precisione. Le pressioni sulle facce opposte di una piastrina sospesa in aria calma, anche microscopica, risultano così praticamente uguali, benché un esperimento opportuno possa rilevare un eccesso di pressione su un lato dell'ordine di un millesimo per cento.

§5. Processi deterministici e processi casuali

Il principio del legame causale fra i fenomeni trova la sua espressione matematica più semplice nello studio dei processi fisici mediante equazioni differenziali, come mostrano gli esempi del §1 del capitolo V.

Sia lo stato del sistema studiato definito, all'istante t, da n parametri

$$x_1 , x_2 , \cdots, x_n .$$

Le loro velocità di variazione si esprimono mediante le derivate rispetto al tempo:

$$\dot{x}_k = \frac{dx_k}{dt} .$$

Se si suppone che queste velocità siano funzioni dei valori dei parametri, si ottiene un sistema di equazioni differenziali:

$$\begin{aligned}
\dot{x}_1 &= f_1(x_1 , x_2 , \cdots, x_n),\\
\dot{x}_2 &= f_2(x_1 , x_2 , \cdots, x_n),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\dot{x}_n &= f_n(x_1 , x_2 , \cdots, x_n).
\end{aligned}$$

La maggior parte delle leggi naturali scoperte alla nascita della fisica matematica, a cominciare dalla legge di Galileo sulla caduta dei gravi, si esprime proprio così. Galileo non

poteva formulare la sua scoperta in questa forma, perché i concetti matematici necessari non erano ancora stati sviluppati; il primo a farlo fu Newton.

In meccanica e in molti altri campi della fisica si esprimono abitualmente queste leggi con equazioni differenziali del secondo ordine. Ciò non introduce però alcun principio nuovo: indicando le velocità $\dot{x}_k$ con i nuovi simboli

$$v_k = \dot{x}_k ,$$

si ottengono, per le derivate seconde delle quantità x_k, le espressioni

$$\frac{d^2x_k}{dt^2} = \dot{v}_k ,$$

e le equazioni del secondo ordine per le n quantità $x_1, \ldots, x_n$ diventano equazioni del primo ordine per le $2n$ quantità $x_1, \ldots, x_n, v_1, \ldots, v_n$.

Consideriamo, per esempio, la caduta di un corpo pesante nell'atmosfera terrestre. Se ci limitiamo a quote modeste, possiamo supporre che la resistenza del mezzo dipenda solo dalla velocità e non dall'altezza. Lo stato del sistema è caratterizzato da due parametri: la distanza z dalla superficie terrestre e la velocità di caduta v, positiva verso il basso. La loro variazione nel tempo è definita dalle due equazioni differenziali

$$\begin{aligned} \dot{z} &= -v, \\ \dot{v} &= g - f(v), \end{aligned} \qquad (31)$$

dove g è l'accelerazione di gravità e $f(v)$ rappresenta una «legge di resistenza» del corpo, espressa come decelerazione.

Se la velocità è modesta e il corpo sufficientemente massiccio, per esempio una pietra di medie dimensioni che cade da pochi metri, la resistenza dell'aria si può trascurare e le (31) diventano

$$\begin{aligned} \dot{z} &= -v, \\ \dot{v} &= g. \end{aligned} \qquad (32)$$

Se all'istante iniziale t_0 le quantità z e v valgono z_0 e v_0, è facile risolvere le (32) e ottenere la formula

$$z = z_0 - v_0(t - t_0) - (g/2)(t - t_0)^2,$$

che descrive l'intera caduta. Per esempio, se $t_0 = 0$ e $v_0 = 0$, si ha

$$z = z_0 - \frac{gt^2}{2} ,$$

la legge trovata da Galileo.

Nel caso generale l'integrazione delle (31) è più difficile, ma, sotto opportune e assai generali condizioni di regolarità su $f(v)$, il risultato fondamentale resta lo stesso: assegnati z_0 e v_0 all'istante t_0, i valori successivi di z e v sono determinati univocamente, fino all'urto con la superficie terrestre. Teoricamente si può eliminare anche quest'ultima limitazione, prolungando la caduta a valori negativi di z. Per il problema così impostato, se f è continua, strettamente crescente con v, parte da $f(0) = 0$ e tende all'infinito per $v \to \infty$, si stabilisce che, proseguendo indefinitamente la caduta, la velocità v tende a un valore costante c, soluzione dell'equazione

$$g = f(c).$$

Il risultato è intuitivamente comprensibile: la velocità di caduta aumenta finché l'effetto accelerante della gravità viene equilibrato dalla resistenza dell'aria. In un lancio con paracadute aperto si raggiunge piuttosto rapidamente una velocità stazionaria, indicata qui come circa cinque metri al secondo.[12] Con paracadute chiuso la resistenza è minore: la velocità stazionaria è quindi maggiore e viene raggiunta solo dopo una caduta molto più lunga.

Per corpi leggeri, come una piuma o un fiocco di lanugine, la fase iniziale di accelerazione è brevissima e spesso inosservabile. La velocità stazionaria si stabilisce molto rapidamente e, con la consueta approssimazione, si può porre $v = c$ durante tutta la caduta. Rimane allora una sola equazione differenziale,

$$\dot{z} = -c,$$

che si integra immediatamente:

$$z = z_0 - c(t - t_0).$$

Così cade un fiocco di lanugine in aria perfettamente calma.

Questa concezione deterministica è trattata in piena generalità nella teoria contemporanea dei sistemi dinamici, cui sono dedicati importanti lavori di matematici sovietici, fra i quali N. N. Bogoljubov e V. V. Stepanov. La teoria comprende anche formulazioni di fenomeni fisici nei quali lo stato del sistema è definito da una o più funzioni, anziché da un numero finito di parametri: per esempio nella meccanica dei mezzi continui. Le leggi elementari di variazione in intervalli di tempo «infinitamente piccoli» sono allora espresse da equazioni alle derivate parziali o con altri mezzi, anziché da equazioni differenziali ordinarie. Due caratteristiche accomunano però tutte le formulazioni deterministiche di processi reali: lo stato del sistema è considerato interamente definito da un ente matematico ω, per esempio un insieme di n numeri reali o una o più funzioni; inoltre il valore ω_0 all'istante iniziale t_0 determina univocamente i valori successivi, per $t > t_0$:

$$\omega = F(t_0, \omega_0, t).$$

Per fenomeni descritti da equazioni differenziali, trovare la funzione F, indicata anche con ϕ, significa, come abbiamo visto, integrare tali equazioni con le condizioni iniziali $\omega = \omega_0$ per $t = t_0$.

I sostenitori del materialismo meccanicistico ritenevano questa formulazione un'espressione esatta e diretta del carattere determinato dei fenomeni reali, cioè del principio fisico di causalità. Secondo Laplace, lo stato del mondo a un dato istante è definito da infiniti parametri, soggetti a infinite equazioni differenziali. Se una «mente universale» potesse scrivere e integrare tutte queste equazioni, potrebbe prevedere con assoluta esattezza l'intera evoluzione del mondo nell'infinito futuro.

Ma questa infinità quantitativa matematica è in realtà assai grossolana rispetto al carattere qualitativamente inesauribile del mondo reale. Né introdurre infiniti parametri, né descrivere lo stato dei mezzi continui mediante funzioni del punto nello spazio basta a rappresentare l'infinita complessità degli eventi effettivi.

Come si è sottolineato nel §3 del capitolo V, lo studio degli eventi reali non procede sempre aumentando il numero dei parametri. In generale è tutt'altro che conveniente complicare l'ente ω che descrive i singoli «stati del sistema» nello schema matematico. L'arte della ricerca consiste piuttosto nel trovare uno spazio Ω molto semplice, cioè un insieme di valori di ω, o di possibili stati diversi del sistema,[13] tale che, sostituendo al processo reale il moto determinato del punto ω in quello spazio, si possano includere tutti gli aspetti essenziali del processo.

Se però astraiamo gli aspetti essenziali di un processo reale, rimane un residuo che dobbiamo considerare casuale. I fattori casuali trascurati esercitano sempre una certa influenza sul processo. Sono pochissimi i fenomeni che, confrontando teoria e osservazione, non mostrino tale influenza. Un caso vicino a questa situazione è il moto dei pianeti sotto la gravità: le loro distanze sono tanto grandi rispetto alle dimensioni che rappresentarli come punti materiali è quasi perfettamente soddisfacente; il materiale nello spazio attraversato è così rarefatto da opporre una resistenza trascurabile; le masse sono tanto grandi che la pressione della luce ha un ruolo quasi nullo. Queste circostanze eccezionali spiegano perché la soluzione matematica per il moto di n punti materiali, i cui stati sono descritti da $6n$ parametri,[14] tenendo conto della sola gravità, concordi in modo così sorprendente con le osservazioni planetarie.

Un caso in parte simile è il volo di un proiettile sotto l'azione della gravità e della resistenza dell'aria. Anche questo è uno dei campi classici in cui i metodi matematici furono relativamente agevoli e ottennero presto grandi successi. Qui però i fattori perturbatori casuali hanno un ruolo assai maggiore: la dispersione, cioè lo scarto dalla traiettoria teorica, raggiunge decine di metri, e alle grandi distanze persino centinaia. Gli scarti dipendono in parte da variazioni casuali della direzione e della velocità iniziali, in parte dalla massa e dal coefficiente di resistenza del proiettile, in parte dalle raffiche, dalle irregolarità del vento e dagli altri fattori delle condizioni straordinariamente complesse e variabili dell'atmosfera terrestre.

La dispersione dei proiettili è studiata dettagliatamente con i metodi della probabilità, e i risultati sono essenziali nella pratica del tiro.

Ma che cosa significa propriamente studiare eventi casuali? Sembrerebbe che, quando il «residuo» casuale di una formulazione diventa troppo grande per essere trascurato, l'unica via sia descrivere il fenomeno più accuratamente, introducendo nuovi parametri e proseguendo uno studio più dettagliato con il medesimo metodo.

In molti casi, però, questo non è realizzabile. Per studiare la caduta di un corpo nell'atmosfera tenendo conto di un vento irregolare e a raffiche, o, come si dice, turbolento, dovremmo sostituire ai due parametri z e v un apparato matematico del tutto ingombrante per descrivere completamente la struttura del moto.

Un procedimento così complesso è necessario solo quando, per qualche ragione, occorre determinare separatamente e in ogni dettaglio l'influenza dei fattori «casuali» residui. Fortunatamente le esigenze pratiche sono di solito diverse: basta stimarne l'effetto complessivo su un lungo intervallo di tempo o su molte ripetizioni del processo.

Consideriamo lo spostamento della sabbia nell'alveo di un fiume o presso un'opera idroelettrica. Di solito la maggior parte della sabbia resta immobile; soltanto ogni tanto una turbolenza particolarmente forte vicino al fondo solleva singoli granelli e li trasporta a una distanza considerevole, depositandoli improvvisamente altrove. In teoria si potrebbe calcolare il moto di ogni granello mediante le leggi dell'idrodinamica. Bisognerebbe però conoscere in ogni dettaglio lo stato iniziale del fondo e della corrente, calcolare passo per passo il flusso, individuare quando la pressione su ciascun granello è sufficiente a metterlo in moto e seguirlo fino all'arresto. È evidente l'assurdità di impostare così un'indagine scientifica concreta. Le leggi medie, o statistiche, dello spostamento della sabbia sul fondo dei fiumi sono invece pienamente accessibili alla ricerca.

Si potrebbero facilmente moltiplicare gli esempi in cui molti fattori casuali producono una legge statistica ben definita. Uno dei più noti e affascinanti, per la vastità delle applicazioni, è la teoria cinetica dei gas: essa mostra come l'influenza congiunta degli urti casuali delle molecole generi leggi precise sulla pressione contro le pareti, sulla diffusione di un gas in un altro, e così via.

§6. Processi casuali di tipo markoviano

Ad A. A. Markov si deve la costruzione di uno schema probabilistico che generalizza direttamente lo schema deterministico del §5, descritto dall'equazione

$$\omega = F(t_0, \omega_0, t).$$

È vero che Markov considerò soltanto uno spazio delle fasi formato da un numero finito di stati, $\Omega = (\omega_1, \omega_2, \ldots, \omega_n)$, e studiò i cambiamenti di stato solo per tempi discreti. Ma in questo modello estremamente schematico riuscì a stabilire una serie di leggi fondamentali.

Al posto della funzione F, che determina univocamente lo stato ω al tempo $t > t_0$ a partire dallo stato ω_0 al tempo t_0, Markov introdusse le probabilità

$$P(t_0, \omega_i; t, \omega_j)$$

di trovarsi nello stato ω_j al tempo t, sapendo che al tempo t_0 il sistema si trovava nello stato ω_i. L'ipotesi markoviana è che, dato lo stato presente, l'ulteriore passato non modifichi le probabilità degli stati futuri. Per tre istanti qualsiasi

$$t_0 < t_1 < t_2$$

queste probabilità sono legate da una relazione introdotta da Markov, che si può chiamare equazione fondamentale di un processo markoviano:

$$\mathbf{P}(t_0, \omega_i ; t_2, \omega_j) = \sum_{k=1}^{n} \mathbf{P}(t_0, \omega_i ; t_1, \omega_k)\mathbf{P}(t_1, \omega_k ; t_2, \omega_j). \qquad (33)$$

Quando lo spazio delle fasi è una varietà continua, il caso più tipico è quello in cui esiste una densità di probabilità $p(t_0, \omega_0; t, \omega)$ per la transizione dallo stato ω_0 allo stato ω nell'intervallo (t_0, t). La probabilità di passare da ω_0 a uno degli stati ω appartenenti a un dominio G dello spazio delle fasi Ω si scrive allora

$$\mathbf{P}(t_0, \omega_0 ; t, G) = \int_G p(t_0, \omega_0 ; t, \omega)\, d\omega, \qquad (34)$$

dove $d\omega$ è un elemento di volume nello spazio delle fasi.[15] Per la densità $p(t_0, \omega_0; t, \omega)$, l'equazione fondamentale (33) assume la forma

$$p(t_0, \omega_0 ; t_2, \omega_2) = \int_\Omega p(t_0, \omega_0 ; t_1, \omega)\, p(t_1, \omega; t_2, \omega_2)\, d\omega. \qquad (35)$$

Risolvere la (35) è generalmente difficile; sotto determinate condizioni, però, se ne possono dedurre equazioni alle derivate parziali più facili da studiare. Alcune furono ricavate dai fisici Fokker e Planck mediante argomenti fisici non rigorosi. Nella sua forma completa, la teoria delle cosiddette equazioni differenziali stocastiche fu sviluppata da autori sovietici: S. N. Bernštejn, A. N. Kolmogorov, I. G. Petrovskij, A. Ja. Hinčin e altri.

Non daremo qui queste equazioni.

Il metodo delle equazioni differenziali stocastiche permette, per esempio, di risolvere agevolmente il problema del moto in aria calma di un corpo piccolissimo, la cui velocità media di caduta c sia molto minore delle velocità caratteristiche degli spostamenti dovuti al «moto browniano». Quest'ultimo nasce perché, date le piccole dimensioni della particella, gli urti delle molecole d'aria sulle sue diverse facce non si equilibrano perfettamente.

Siano c la velocità media di caduta e D il cosiddetto coefficiente di diffusione. Supponiamo che la particella non resti sulla superficie terrestre, $z = 0$, ma venga «riflessa», ossia risospinta nell'atmosfera dalle forze browniane; supponiamo inoltre che all'istante t_0 si trovi all'altezza z_0. La densità di probabilità $p(t_0, z_0; t, z)$ di trovarla all'altezza z all'istante t, per $c > 0$ e $D > 0$, è espressa dalla formula

$$p(to, zo; t, z) = \frac{1}{2\sqrt{(\pi D \Delta t)}} \left[e^{\frac{-(z - zo)^2}{4D\Delta t}} + e^{\frac{-(z + zo)^2}{4D\Delta t}} \right]$$

$$\times e^{\frac{-c(z - zo)}{2D} - \frac{c^2 \Delta t}{4D}}$$

$$+ \frac{c}{D\sqrt{\pi}} e^{-cz/D} \int_{\frac{z + zo - c\Delta t}{2\sqrt{(D\Delta t)}}}^{\infty} e^{-u^2} du,$$

$$\Delta t = t - to > 0.$$

Nella figura 4 mostriamo come possono variare le curve $p(t_0, z_0; t, z)$ per una successione di istanti t.

La posizione della particella, inizialmente nota, diventa indeterminata, «casuale»; la sua altezza media può anche aumentare, come nel caso descritto più avanti con partenza dal suolo. L'aspetto più interessante è che, per qualsiasi t_0 e z_0, al tendere di t all'infinito,

$$p(t_0, z_0; t, z) \to \frac{c}{D} e^{-cz/D}; \qquad (36)$$

Esiste cioè una distribuzione limite dell'altezza della particella, e la speranza matematica dell'altezza tende, al crescere di t, al limite positivo

$$z^* = \frac{c}{D} \int_0^\infty z e^{-cz/D} dz = \frac{D}{c}. \qquad (37)$$

Sebbene dunque la particella, finché si trova sopra il suolo, tenda a cadere per effetto della gravità, nel protrarsi di questo vagare nell'atmosfera essa si trova in media a una determinata altezza positiva. Se l'altezza iniziale z_0 è minore di z^*, dopo un intervallo sufficientemente lungo la posizione media risulta più alta di quella iniziale. È ciò che mostra la figura 5, dove $z_0 = 0$.

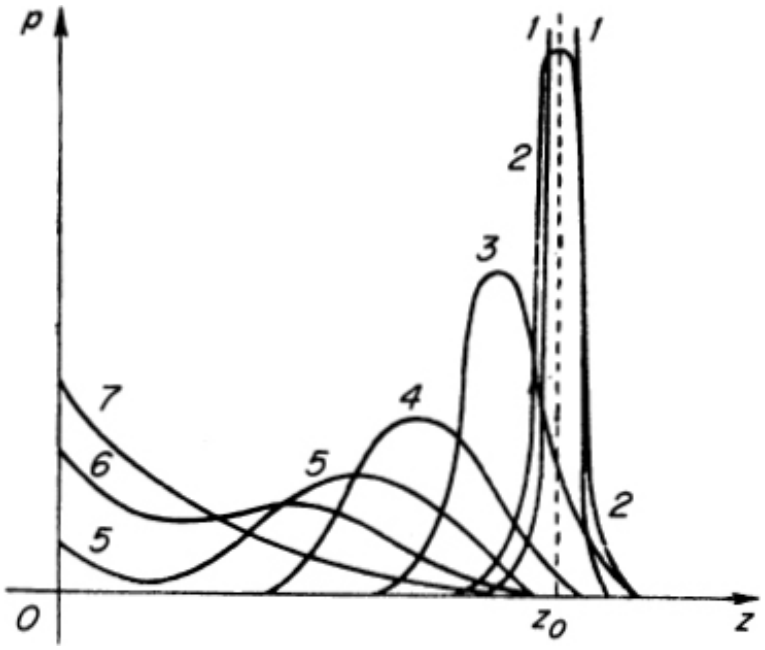


Figura 4.

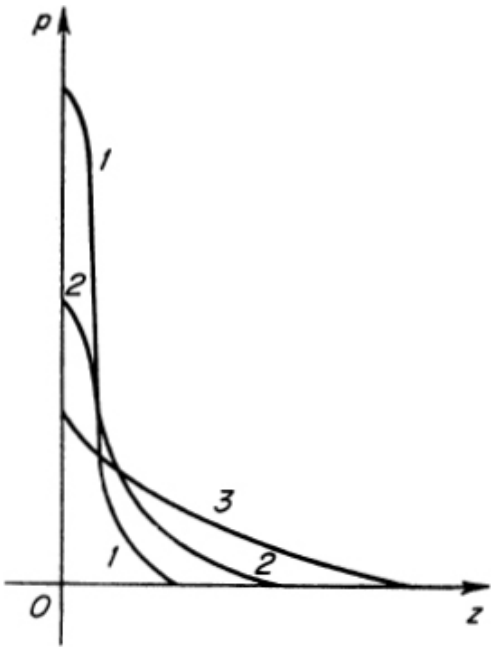


Figura 5.

Per una singola particella, il valore medio z^* è soltanto una speranza matematica; dalla legge dei grandi numeri segue però che, per molte particelle nelle opportune condizioni di indipendenza, esso si realizza nelle medie osservate. La distribuzione delle particelle rispetto

all'altezza segue le leggi indicate e, dopo un tempo sufficiente, la sua densità si stabilizza secondo la (36).

Quanto detto è direttamente applicabile soltanto a gas, fumo e simili presenti nell'aria in piccole concentrazioni, perché le quantità c e D sono state supposte determinate da uno stato dell'atmosfera prefissato. Con alcune complicazioni, tuttavia, la teoria si applica anche alla diffusione reciproca dei gas che compongono l'atmosfera e alla distribuzione delle loro densità con l'altezza.

Il rapporto c/D aumenta con le dimensioni delle particelle, e il carattere del moto passa dalla diffusione alla caduta regolare secondo le leggi del §5. La teoria permette di seguire tutte le transizioni fra il moto puramente diffusivo e queste leggi di caduta.

Il moto delle particelle sospese in un'atmosfera turbolenta è un problema più difficile, ma in linea di principio si può trattare con metodi probabilistici analoghi.

Letture consigliate

H. Cramer, The elements of probability theory and some of its applications, Wiley, New York, 1955.

W. K. Feller, An introduction to probability theory and its applications, 2nd ed., Wiley, New York, 1957.

B. V. Gnedenko and A. Ya. Khinchin, An elementary introduction to the theory of probability, W. H. Freeman and Co., San Francisco, Calif., 1961.

M. Kac, Statistical independence in probability, analysis and number theory, Wiley, New York, 1959.

J. G. Kemeny and J. L. Snell, Finite Markov chains, Van Nostrand, New York, 1960.

A. N. Kolmogorov, Foundations of the theory of probability, Chelsea, New York, 1950.

M. Loève, Probability theory: foundations, random sequences, Van Nostrand, New York, 1955.

E. Parzen, Modern probability theory and its applications, Wiley, New York, 1960.

W. A. Whitworth, Choice and chance, Stechert, New York, 1942.

Note al capitolo XI

[1] Per brevità scriveremo ora P(A) al posto di P(A/S).

[2] Analogamente, l'intersezione C di un numero qualsiasi di eventi A_1, A_2, ..., A_s consiste nel verificarsi di tutti gli eventi assegnati.

[3] Un significato più preciso dell'indipendenza degli esperimenti è il seguente. Dividiamo in un modo qualsiasi gli n esperimenti in due gruppi. Sia A l'evento che tutti gli esperimenti del primo gruppo abbiano esiti prefissati, e B l'evento che quelli del secondo abbiano esiti prefissati. Gli esperimenti si dicono congiuntamente indipendenti se, per ogni suddivisione e ogni scelta degli esiti, A e B sono indipendenti nel senso della (4).

Nel §4 torneremo sul significato oggettivo dell'indipendenza degli eventi nel mondo reale.

[4] Questa ipotesi di indipendenza è confermata dall'esperienza nel modello considerato.

[5] La dimostrazione della stima (8) sarà discussa più avanti in questo paragrafo.

[6] Questa è la notazione usuale per la stima (8) della probabilità che valga la (7).

[7] Qui μ assume i valori $m = 0, 1, \ldots, n$ con probabilità P_m; cioè

$$\mathbf{P}(\mu = m) = P_m .$$

[8] Il coefficiente di correlazione fra $\xi^{(i)}$ e $\xi^{(j)}$ è l'espressione

$$R = \frac{M\{[\xi^{(i)} - M(\xi^{(i)})][\xi^{(j)} - M(\xi^{(j)})]\}}{\sigma_{\xi^{(i)}}\sigma_{\xi^{(j)}}}.$$

Se le deviazioni standard $\sigma_{\xi^{(i)}}$ e $\sigma_{\xi^{(j)}}$ sono positive, la (24) equivale a $R = 0$.

Il coefficiente R caratterizza la dipendenza lineare fra le variabili aleatorie. Vale sempre $|R| \leqslant 1$, e $R = \pm 1$ soltanto quando sussiste, con probabilità uno, una relazione lineare

$$\eta = a\xi + b \qquad (a \neq 0).$$

Per variabili indipendenti con varianze positive, $R = 0$.

[9] Due variabili aleatorie ξ e η, che possono assumere rispettivamente i valori $x_1, \ldots, x_m$ e $y_1, \ldots, y_n$, sono indipendenti se, per ogni i e j, gli eventi $A_i = \{\xi = x_i\}$ e $B_j = \{\eta = y_j\}$ sono indipendenti nel senso definito nel §2.

[10] Lo stesso vale, per ogni base intera $n \geq 2$, per le cifre α_k nello sviluppo di ξ

$$\xi = \alpha_1/n + \alpha_2/n^2 + \alpha_3/n^3 + \ldots .$$

[11] Per la natura dei suoi problemi, la teoria della probabilità resta tuttavia una disciplina matematica autonoma: i suoi risultati fondamentali, presentati dettagliatamente nel §3, appaiono artificiali e non necessari dal punto di vista della pura teoria della misura.

[12] L'affermazione va intesa nel senso che, in pratica, v si avvicina rapidamente a c.

[13] Nell'esempio precedente del corpo in caduta, lo spazio delle fasi è l'insieme delle coppie (z, v), cioè un piano. Per gli spazi delle fasi in generale, si vedano i capitoli XVII e XVIII.

[14] Le tre coordinate e le tre componenti della velocità di ciascun punto.

[15] Propriamente, la (34) serve a definire la densità di probabilità. La quantità $p\, d\omega$ è uguale, a meno di un infinitesimo di ordine superiore, alla probabilità di passare, nel tempo da t_0 a t, dallo stato ω_0 all'elemento di volume $d\omega$.

Nota editoriale

La traduzione conserva le sei sezioni, le cinque figure, le formule, la tavola numerica, la bibliografia e le quindici note. Le parole inglesi incorporate nelle formule sono tradotte. Sono corretti refusi nella formula binomiale per 40 prove, nei simboli P_r e σ_ξ, negli indici degli stati di Markov e nello sviluppo in base n. Nella legge della caduta compare la velocità iniziale v_0 e il termine gravitazionale è $-(g/2)(t - t_0)^2$. Nella densità del processo riflesso, ricomposta, il fattore esponenziale contiene $c(z - z_0)$ e il denominatore dell'ultimo termine contiene $\sqrt{\pi}$; la variabile muta dell'integrale è u.

Le uguaglianze numeriche approssimate sono distinte dalle uguaglianze esatte. La tavola di F conserva le cifre dell'originale; nella (19) F(4) è arrotondato a 0,99994. Nell'esempio binomiale con n = 100, p = 0,5 e μ > 65, il calcolo esatto dà circa 0,000895: il valore 0,0015 dell'originale è sostituito. I criteri di significatività e le valutazioni sulla ricerca scientifica sono riportati nel loro contesto storico; gli esempi hanno funzione illustrativa.

Si esplicitano le condizioni di positività nei quozienti e nelle deviazioni standard, l'indipendenza delle prove nei calcoli binomiali e il limite uniforme delle varianze nella legge dei grandi numeri. L'indipendenza e l'uguaglianza delle varianze, da sole, non garantiscono il teorema centrale del limite per variabili non identicamente distribuite: le formule sono intese sotto le ulteriori ipotesi del teorema pertinente. Il coefficiente di correlazione descrive la dipendenza lineare; correlazione nulla non implica in generale indipendenza. L'ordine $n \square 1/\varepsilon^2$ indica una scala di grandezza, non un numero universale di osservazioni sufficiente a ogni livello di fiducia.

Per l'esempio termodinamico si intende acqua pura in equilibrio e si distingue lo stato macroscopico dalla presenza di vapore. Nelle equazioni della caduta v è positiva verso il basso e f è una decelerazione; sono rese esplicite condizioni sufficienti per la velocità limite. I richiami al materialismo e alla causalità conservano il punto di vista dell'autore. La proprietà markoviana è esplicitata nel testo; per la costruzione generale della probabilità come misura occorre l'additività numerabile, oltre agli assiomi elementari enunciati.

Il termine storico «equazioni differenziali stocastiche» nel §6 si riferisce qui soprattutto alle equazioni per le probabilità di transizione, e non soltanto al significato moderno di equazione differenziale guidata da rumore. Il modello browniano descritto ha un confine riflettente in z = 0 e coefficienti c, D positivi e costanti. La posizione diventa casuale, ma l'altezza media non cresce per ogni altezza iniziale: il testo precisa il caso di partenza dal suolo. Le medie osservate su molte particelle richiedono adeguate ipotesi di indipendenza o debole dipendenza. Le affermazioni storiche e sui risultati disponibili restano riferite all'epoca dell'opera.

CAPITOLO XII

L’approssimazione delle funzioni

S. M. Nikol’skij

§1. Introduzione

Nella vita pratica affrontiamo continuamente il problema di approssimare certi numeri mediante altri. Le misure di grandezze concrete, come lunghezze, aree e temperature, ci conducono a numeri che sono soltanto approssimati. Nella pratica usiamo numeri razionali, cioè della forma p/q, con p e q interi e $q \neq 0$. Esistono però anche i numeri irrazionali: benché non li usiamo direttamente nelle misurazioni, i ragionamenti teorici ci conducono spesso a essi. Sappiamo, per esempio, che la lunghezza di una circonferenza di raggio $r = \frac{1}{2}$ è il numero irrazionale π, e che l’ipotenusa di un triangolo rettangolo con cateti unitari ha lunghezza $\sqrt{2}$. Nei calcoli effettivi si approssimano anzitutto gli irrazionali mediante razionali con la precisione richiesta, di solito mediante frazioni decimali finite.

Lo stesso accade per le funzioni. Le leggi quantitative della natura sono espresse matematicamente da funzioni con vari gradi di precisione, non con esattezza assoluta. Inoltre, anche per funzioni definite da regole interamente matematiche, in moltissimi casi occorre approssimarle mediante altre funzioni con precisione assegnata, per poterle calcolare concretamente.

Queste osservazioni non riguardano soltanto il calcolo. Definire una funzione mediante altre funzioni ha grande importanza teorica. L’analisi matematica ha portato a scoprire e studiare classi molto importanti di funzioni approssimanti che, sotto determinate condizioni, costituiscono strumenti naturali per approssimare funzioni più o meno arbitrarie. Fra queste classi vi sono anzitutto i polinomi algebrici e trigonometrici e le loro generalizzazioni. Si è mostrato che dalle proprietà della funzione da approssimare si può stimare, in opportune condizioni, il suo scarto da una determinata successione di funzioni approssimanti. Viceversa, conoscendo il comportamento di questo scarto si possono stabilire proprietà della funzione. Si è così sviluppata una teoria delle funzioni fondata sulla loro rappresentazione approssimata mediante diverse classi di approssimanti. Una teoria analoga esiste nella teoria dei numeri: le proprietà degli irrazionali vengono studiate attraverso le loro approssimazioni razionali.

Nel capitolo II il lettore ha già incontrato un metodo fondamentale di approssimazione: la formula di Taylor. Essa permette, sotto certe condizioni, di approssimare una funzione mediante un’altra della forma $P(x) = a_0 + a_1 x + \ldots + a_n x^n$, detta polinomio algebrico. I coefficienti a_k sono costanti, indipendenti da x.

Un polinomio algebrico ha una struttura molto semplice: per calcolarlo, dati i coefficienti e il valore di x, bastano addizione, sottrazione e moltiplicazione. Questa semplicità è importantissima nella pratica ed è una delle ragioni per cui i polinomi algebrici sono lo

strumento più diffuso per approssimare funzioni; di un'altra ragione parleremo più avanti. Basti ricordare che i calcoli tecnici si eseguono ormai su vasta scala mediante macchine calcolatrici. Esse lavorano rapidamente e senza stancarsi, ma possono compiere soltanto operazioni relativamente semplici. Si possono programmare per operazioni aritmetiche su numeri molto grandi, ma non per eseguire, per esempio, un processo infinito di passaggio al limite. Una macchina non calcola in generale log x esattamente; possiamo però approssimarlo, su un intervallo opportuno, con un polinomio $P(x)$ alla precisione richiesta e poi calcolare quest'ultimo con la macchina.

Oltre alla formula di Taylor, altre formule hanno grande importanza pratica nell'approssimazione mediante polinomi algebrici. Fra esse vi sono le formule d'interpolazione, usate in particolare nel calcolo approssimato degli integrali e nell'integrazione numerica delle equazioni differenziali. È ben noto anche il metodo di approssimazione in media quadratica, ampiamente impiegato con funzioni diverse dai polinomi algebrici. Per alcune questioni pratiche ha grande importanza la migliore approssimazione uniforme, o di Čebyšev, introdotta dal grande matematico russo. Come vedremo, il metodo nacque dalla soluzione di un problema di costruzione dei meccanismi.

Vogliamo dare al lettore un'idea di questi metodi e, per quanto possibile, indicare le condizioni in cui uno è preferibile agli altri. Nessuno è il migliore in assoluto: ciascuno può risultare superiore sotto determinate condizioni. In un problema fisico, per esempio, il metodo più adatto ad approssimare le funzioni coinvolte è suggerito dalla natura del problema, ossia da considerazioni fisiche. Vedremo anche che, in certe condizioni, un metodo è applicabile mentre un altro non lo è.

Ogni metodo di calcolo è nato in un'epoca propria e possiede una teoria e una storia caratteristiche. Newton conosceva già una formula d'interpolazione e le diede una forma molto comoda per il calcolo mediante le cosiddette differenze divise. Il metodo dell'approssimazione in media quadratica risale, all'epoca dell'opera, ad almeno centocinquant'anni prima. Per molto tempo, tuttavia, questi metodi non formarono una teoria organica: erano procedimenti pratici distinti, e non erano chiari neppure i limiti della loro applicabilità.

La teoria moderna dell'approssimazione delle funzioni nacque dai lavori di Čebyšev. Egli introdusse il concetto fondamentale di migliore approssimazione, in particolare uniforme, lo applicò sistematicamente a problemi pratici e ne sviluppò le basi teoriche. Dopo di lui, le sue idee furono sviluppate dagli allievi E. I. Zolotarev, A. N. Korkin e dai fratelli A. A. e V. A. Markov. In questa fase furono introdotti i concetti fondamentali e trovati metodi, ancora largamente usati, per ottenere le migliori approssimazioni di singole funzioni arbitrarie. Furono inoltre studiate le proprietà delle classi approssimanti, soprattutto dei polinomi algebrici e trigonometrici, in rapporto alle esigenze dei problemi pratici.

Lo sviluppo successivo fu influenzato da un'importante scoperta del matematico tedesco Weierstrass, alla fine dell'Ottocento: egli dimostrò rigorosamente che ogni funzione continua su un intervallo chiuso e limitato può essere approssimata da un polinomio algebrico con qualsiasi precisione assegnata. Questa è la seconda ragione del carattere universale dei

polinomi come strumenti d'approssimazione. La semplicità della loro costruzione non basta: occorre anche poter approssimare ogni funzione continua con un errore arbitrariamente piccolo. Weierstrass dimostrò precisamente questa possibilità.

Le idee di Čebyšev e il teorema di Weierstrass furono alla base, all'inizio del Novecento, dello sviluppo contemporaneo della teoria. Ricordiamo i nomi di S. N. Bernštejn, Borel, Jackson, Lebesgue e de la Vallée-Poussin. In breve, mentre nella fase di Čebyšev e fino all'inizio del Novecento si studiava soprattutto l'approssimazione di singole funzioni, il problema caratteristico della fase successiva è l'approssimazione, mediante polinomi o altri strumenti, di intere classi di funzioni: analitiche, differenziabili e così via.

La scuola russa, poi sovietica, ha avuto un ruolo di primo piano. Contributi importanti sono dovuti a S. N. Bernštejn, A. N. Kolmogorov, M. A. Lavrent'ev e ai loro allievi. La teoria è ormai diventata un ramo sostanzialmente autonomo della teoria delle funzioni.

Oltre ai polinomi algebrici, uno strumento importantissimo di approssimazione è costituito dai polinomi trigonometrici. Un polinomio trigonometrico di ordine n è una funzione della forma

$$u_n(x) = \alpha_0 + \alpha_1 \cos x + \beta_1 \sin x + \alpha_2 \cos 2x + \beta_2 \sin 2x + \cdots + \alpha_n \cos nx + \beta_n \sin nx,$$

oppure, più sinteticamente,

$$u_n(x) = \alpha_0 + \sum_{k=1}^{n} (\alpha_k \cos kx + \beta_k \sin kx),$$

dove α_k e β_k sono costanti.

I diversi metodi di approssimazione mediante polinomi trigonometrici sono generalmente collegati in modo piuttosto semplice ai corrispondenti metodi per i polinomi algebrici. Fra essi ha un ruolo particolare lo sviluppo in serie di Fourier, di cui parleremo nel §7. Queste serie prendono il nome dal matematico francese che, all'inizio dell'Ottocento, ottenne risultati teorici su di esse studiando la conduzione del calore. Le serie trigonometriche erano però già state studiate a metà del Settecento da Leonhard Euler e Daniel Bernoulli: per Eulero erano legate alle ricerche astronomiche, per Bernoulli alla corda vibrante. Essi posero il problema fondamentale della rappresentabilità di una funzione più o meno arbitraria mediante una serie trigonometrica. La risposta, precisata solo nell'Ottocento e discussa più avanti, fu anticipata nelle sue linee essenziali da Bernoulli.

Le serie di Fourier hanno grande importanza in fisica, ma daremo poco spazio a questo aspetto, già trattato nel capitolo VI. Lì il lettore troverà anche problemi fisici che conducono naturalmente a sviluppi in serie diverse da quelle trigonometriche, ma a esse molto simili: le cosiddette serie di funzioni ortogonali.

Le serie di Fourier hanno, all'epoca del testo, una storia di due secoli. Non sorprende quindi che la loro teoria sia divenuta vastissima, sottile e profonda, fino a costituire una disciplina

autonoma. Un ruolo particolarmente notevole vi ha avuto la scuola moscovita della teoria delle funzioni di variabile reale: N. N. Luzin, A. N. Kolmogorov, D. E. Men'šov e altri.

L'importanza dei polinomi trigonometrici nella matematica contemporanea non si esaurisce nell'approssimazione. Nel capitolo X, per esempio, il lettore ha incontrato i risultati fondamentali di I. M. Vinogradov nella teoria dei numeri, ottenuti mediante un opportuno apparato di somme trigonometriche.

§2. I polinomi d'interpolazione

Un caso particolare di costruzione del polinomio interpolatore.

Nei calcoli pratici il metodo d'interpolazione è molto usato. Per introdurre queste questioni consideriamo il seguente problema elementare.

Sia data sull'intervallo $[x_0, x_2]$ la funzione $y = f(x)$, con il grafico della figura 1. Il suo aspetto ricorda un arco di parabola. Per approssimarla con una funzione semplice è quindi naturale scegliere un polinomio di secondo grado,

$$P(x) = a_0 + a_1 x + a_2 x^2, \qquad (1)$$

il cui grafico è una parabola.

Il metodo consiste nel scegliere un punto interno x_1 all'intervallo $[x_0, x_2]$. Ai tre punti x_0, x_1, x_2 corrispondono i valori della funzione

$$y_0 = f(x_0),\ y_1 = f(x_1),\ y_2 = f(x_2).$$

Costruiamo un polinomio (1) che coincida con la funzione in questi tre punti; il suo grafico è tratteggiato nella figura 1. Dobbiamo cioè scegliere i coefficienti a_0, a_1, a_2 in modo che soddisfino le equazioni

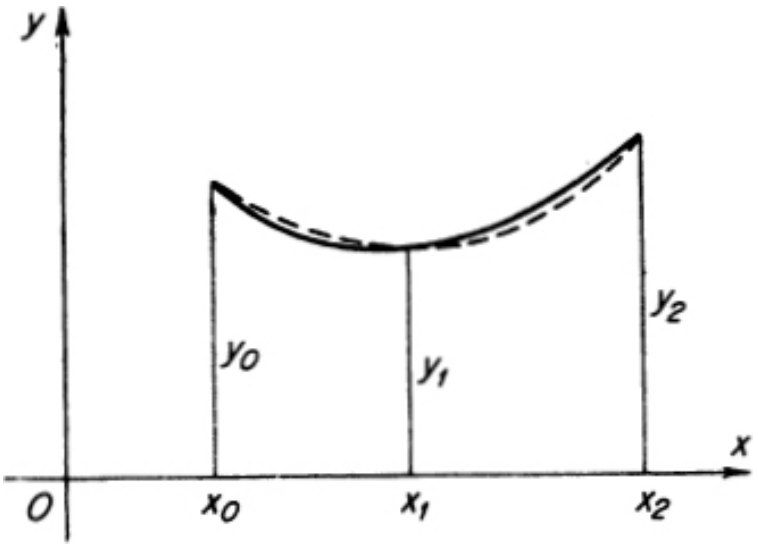


Figura 1.

$$P(x_0) = y_0\,,\ P(x_1) = y_1\,,$$

$$P(x_2) = y_2\,. \qquad (2)$$

La funzione $f(x)$ non deve necessariamente essere definita da una formula: i suoi valori possono, per esempio, essere dati empiricamente, come nel grafico della figura 1. Per risolvere il problema scegliamo l'approssimante come espressione analitica, precisamente il polinomio $P(x)$. Se la precisione è soddisfacente, esso offre il vantaggio di permettere il calcolo dei valori intermedi.

Potremmo risolvere il problema impostando le tre equazioni

$$y_0 = a_0 + a_1x_0 + a_2x_0^2\,,$$

$$y_1 = a_0 + a_1x_1 + a_2x_1^2\,,$$

$$y_2 = a_0 + a_1x_2 + a_2x_2^2\,,$$

risolvendole rispetto ad a_0, a_1, a_2 e sostituendo i risultati nella (1). Procediamo però diversamente. Costruiamo anzitutto il polinomio $Q_0(x)$ di secondo grado che soddisfa $Q_0(x_0) = 1$, $Q_0(x_1) = 0$, $Q_0(x_2) = 0$. Dalle ultime due condizioni segue che deve avere la forma $A(x - x_1)(x - x_2)$; dalla prima,

$$A = \frac{1}{(x_0 - x_1)(x_0 - x_2)}\,.$$

Il polinomio cercato è quindi

$$Q_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}\,.$$

Analogamente, i polinomi

$$Q_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}\,, \qquad Q_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

soddisfano le condizioni

$$Q_1(x_0) = Q_1(x_2) = 0, \quad Q_1(x_1) = 1,$$
$$Q_2(x_0) = Q_2(x_1) = 0, \quad Q_2(x_2) = 1.$$

È evidente che $y_0Q_0(x)$ vale y_0 per $x = x_0$ e si annulla per $x = x_1$ e $x = x_2$. Proprietà corrispondenti valgono per $y_1Q_1(x)$ e $y_2Q_2(x)$.

Ne segue subito che il polinomio interpolatore cercato è dato dalla formula

$$P(x) = y_0 Q_0(x) + y_1 Q_1(x) + y_2 Q_2(x)$$
$$= y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}$$
$$+ y_2 \frac{(x - x_0)(x - x_1)}{(x_2 + x_0)(x_2 - x_1)}. \qquad (3)$$

Esso è l'unico polinomio di grado non superiore a due che risolve il problema. Se infatti anche $P_1(x)$ lo risolvesse, la differenza $P_1(x) - P(x)$, di grado non superiore a due, si annullerebbe nei tre punti x_0, x_1, x_2. Dall'algebra sappiamo che un tale polinomio con tre zeri distinti è identicamente nullo. Dunque i due polinomi coincidono.

In generale il polinomio ottenuto coincide con la funzione data nei tre nodi e ne differisce negli altri punti.

Se x_1 è il punto medio dell'intervallo e poniamo $x_2 - x_1 = x_1 - x_0 = h$, la (3) si semplifica:

$$P(x) = \frac{1}{2h^2} [y_0(x - x_1)(x - x_2) - 2y_1(x - x_0)(x - x_2) + y_2(x - x_0)(x - x_1)].$$

Interpoliamo, per esempio, la sinusoide $y = \sin x$, figura 2, con un polinomio quadratico che coincida con essa per $x = 0, \pi/2, \pi$. Evidentemente esso ha la forma

$$P(x) = \frac{4}{\pi^2} x(\pi - x) \approx \sin x.$$

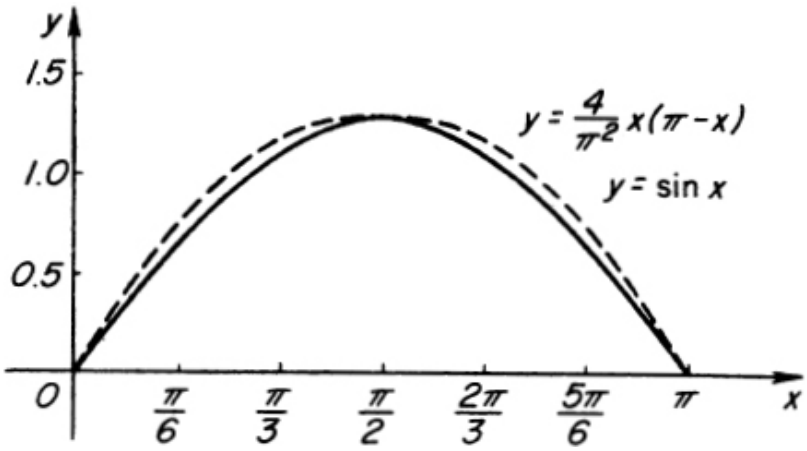


Figura 2.

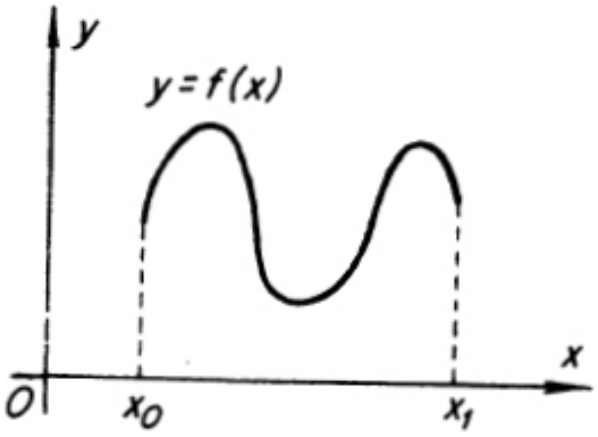


Figura 3.

Confrontiamo sin x e $P(x)$ in due punti intermedi:

$$P(\pi/4) = 0{,}75, \text{ mentre } \sin(\pi/4) = \sqrt{2}/2 \approx 0{,}71;$$
$$P(\pi/6) = 10/18, \text{ mentre } \sin(\pi/6) = 9/18.$$

Abbiamo così approssimato sin x sull'intervallo $[0, \pi]$ con una precisione[1] di circa 0,06. Lo sviluppo di Taylor intorno a $\pi/2$ dà invece

$$\sin x = \cos\left(\frac{\pi}{2} - x\right) = 1 - \frac{\left(\frac{\pi}{2} - x\right)^2}{2!} + \frac{\left(\frac{\pi}{2} - x\right)^4}{4!} - \cdots.$$

Fermandoci al secondo termine, per $x = 0$ otteniamo l'approssimante $1 - \pi^2/8 \approx -0{,}234$ per sin 0, ossia un errore maggiore di 0,2.

L'interpolazione ha quindi prodotto, sull'intero intervallo $[0, \pi]$, un'approssimazione quadratica migliore di quella di Taylor del secondo grado. Non bisogna però dimenticare che, vicino al centro $x = \pi/2$ dello sviluppo, la formula di Taylor è molto precisa, più dell'interpolazione ottenuta.

La soluzione generale del problema.

Una funzione più complessa, come quella della figura 3, difficilmente si presta a essere approssimata da un polinomio quadratico: nessuna parabola può seguirne tutte le curve. È naturale tentare un'interpolazione con un polinomio di grado superiore, nell'esempio almeno quarto.

Il problema generale consiste nel costruire un polinomio $P(x) = a_0 + a_1x + \ldots + a_nx^n$, di grado non superiore a n, che coincida con la funzione in $n + 1$ punti distinti, cioè soddisfi le equazioni

$$P(x_0) = f(x_0),\ P(x_1) = f(x_1),\ \cdots,\ P(x_n) = f(x_n).$$

I punti in cui si impone la coincidenza fra funzione e polinomio si chiamano nodi d'interpolazione.

Ragionando come nel caso quadratico, si dimostra facilmente che il polinomio cercato si scrive nella forma

$$P_n(x) = \sum_{k=0}^{n} \frac{(x - x_0)(x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n)}{(x_k - x_0)(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_n)} f(x_k), \tag{4}$$

e che è unico fra i polinomi di grado non superiore a n. Questa è la formula di Lagrange. Essa può assumere altre forme: molto usata nei calcoli è quella di Newton mediante differenze divise.

Lo scarto fra polinomio interpolatore e funzione.

L'interpolazione è uno strumento generale di approssimazione. Per costruire l'interpolatore non si richiedono proprietà particolari della funzione, per esempio l'esistenza di derivate sull'intero intervallo: sotto questo aspetto il metodo ha un vantaggio sulla formula di Taylor. Esistono inoltre funzioni analitiche in ogni punto di un intervallo che non possono essere approssimate su tutto l'intervallo mediante un unico sviluppo di Taylor prefissato. Supponiamo di voler approssimare $1/(1 + x^2)$ su $[-2, 2]$ con un polinomio. Si potrebbe provare lo sviluppo intorno a $x = 0$:

$$1/(1 + x^2) = 1 - x^2 + x^4 - x^6 + \ldots, \quad |x| < 1.$$

Ma la serie converge soltanto per $-1 < x < 1$. Fuori da $[-1, 1]$ diverge, e quindi non può approssimare la funzione sull'intero intervallo $[-2, 2]$. Il metodo d'interpolazione resta invece applicabile, con un'opportuna scelta dei nodi.

Naturalmente, in ogni caso occorre scegliere numero e distribuzione dei nodi in modo che l'errore soddisfi le richieste. Per funzioni con derivate di ordine abbastanza alto, una risposta sulla grandezza possibile dell'errore è data dal seguente risultato classico, che enunciamo senza dimostrazione.

Se $f(x)$ ha derivata continua di ordine $n + 1$ su $[x_0, x_n]$, allora, per ogni punto intermedio x, lo scarto dal polinomio di Lagrange con nodi $x_0 < x_1 < \ldots < x_n$ è dato da

$$f(x) - P(x) = [(x - x_0)(x - x_1) \cdots (x - x_n)/(n + 1)!] f^{(n+1)}(c),$$

dove c è un punto fra x_0 e x_n, in generale dipendente da x. La formula ricorda il resto di Taylor e ne è sostanzialmente una generalizzazione. Se la derivata $f^{(n+1)}$ non supera in valore assoluto M sull'intervallo, l'errore soddisfa, per ogni x, la stima

$$|f(x) - P_n(x)| \leq [|x - x_0| \cdots |x - x_n|/(n + 1)!] M.$$

La teoria contemporanea offre molti altri modi di stimare l'errore d'interpolazione. Lo studio approfondito del problema ha rivelato fatti interessanti e del tutto inattesi.

Consideriamo una funzione regolare $y = f(x)$ su $[-1, 1]$, cioè con derivata prima continua, il cui grafico sia una curva continua con tangente variabile continuamente. La scelta degli estremi -1 e 1 non è essenziale: con modifiche irrilevanti, quanto segue vale su qualsiasi intervallo $[a, b]$.

Supponiamo di aver scelto su $[-1, 1]$ un sistema di $n + 1$ punti

$$-1 \leqslant x_0 < x_1 < \cdots < x_n \leqslant 1 \qquad (5)$$

Costruiamo poi il polinomio $P(x) = a_0 + a_1 x + \cdots + a_n x^n$, di grado non superiore a n, che coincide con $f(x)$ in questi punti. Supponiamo per il momento che i punti del sistema (5) siano equidistanti nell'intervallo. Al crescere indefinito di n, il corrispondente polinomio interpolatore $P_n(x)$ coinciderà con $f(x)$ in un numero sempre maggiore di punti. Si potrebbe quindi pensare che, anche in un punto intermedio x, non appartenente al sistema (5), la differenza $f(x) - P_n(x)$ tenda a zero per $n \to \infty$. Questa opinione era ancora diffusa alla fine dell'Ottocento, ma in seguito si scoprì che le cose stanno ben diversamente. Si è dimostrato che, per molte funzioni regolari, e persino analitiche, quando i nodi x_k sono equidistanti, i polinomi interpolatori $P_n(x)$ non convergono affatto a $f(x)$ per $n \to \infty$. Il grafico del polinomio passa certamente per i punti d'interpolazione assegnati; tuttavia, per grandi valori di n, fra questi punti può discostarsi fortemente dal grafico della funzione, e lo scostamento può crescere al crescere di n. Ulteriori ricerche hanno mostrato che questo inconveniente si può evitare, almeno per le funzioni regolari, distribuendo i nodi più radi al centro dell'intervallo e più fitti verso gli estremi. Si è infatti dimostrato che, in un senso ben preciso, la distribuzione migliore si ottiene prendendo come nodi x_k gli zeri[2] del polinomio di Čebyšev $\cos[(n + 1) \arccos x]$, dati dalla formula

$$x_k = \cos \frac{2k + 1}{2(n + 1)} \pi \qquad (k = 0, 1, \cdots, n).$$

I polinomi interpolatori corrispondenti a questi nodi di Čebyšev convergono uniformemente alla funzione da cui sono costruiti, purché questa sia regolare, cioè continua e dotata di derivata prima continua. Il grafico di una funzione siffatta è una curva continua la cui tangente varia con continuità. La figura 4 mostra la distribuzione degli zeri del polinomio di Čebyšev nel caso $n = 5$.

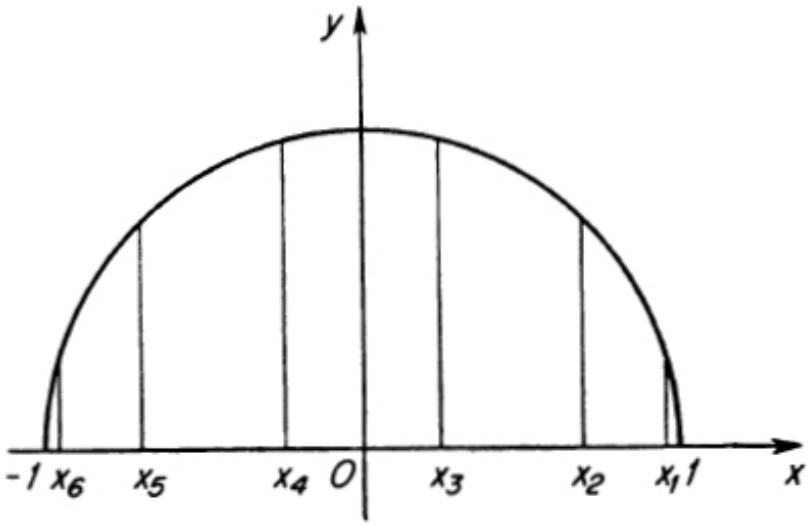


Figura 4.

Per le funzioni continue arbitrarie, senza ulteriori ipotesi di regolarità, la situazione è peggiore: non esiste uno schema prefissato di nodi d'interpolazione che renda il procedimento convergente per ogni funzione continua; questo è il teorema di Faber. In altri termini, comunque si scelgano nell'intervallo [−1, 1] sistemi di nodi il cui numero tende all'infinito, si può trovare una funzione $f(x)$, continua nell'intervallo, per la quale la successione dei corrispondenti polinomi interpolatori non converge alla funzione. Se fosse stato noto, questo fatto sarebbe parso paradossale anche ai matematici della metà dell'Ottocento. La spiegazione è che fra le funzioni continue non regolari ve ne sono di straordinariamente «cattive», per esempio funzioni prive di derivata in ogni punto dell'intervallo di definizione: esse mostrano quanto possa essere ampia la classe in cui cercare controesempi a un dato procedimento d'interpolazione. Modificando il procedimento precedente si possono costruire metodi efficaci di approssimazione polinomiale anche per queste funzioni, ma non ci soffermeremo qui su tale questione.

Osserviamo infine che i polinomi algebrici non sono l'unico strumento d'interpolazione. Esistono, per esempio, metodi d'interpolazione mediante polinomi trigonometrici, ampiamente sviluppati sia sul piano pratico sia su quello teorico.

§3. Approssimazione degli integrali definiti

L'interpolazione delle funzioni trova numerose applicazioni nel calcolo approssimato degli integrali. Come esempio ricaveremo una formula per l'approssimazione di un integrale definito: la formula di Simpson, largamente usata nell'analisi applicata.

Si voglia calcolare approssimativamente l'integrale definito, sull'intervallo $[a, b]$, della funzione $f(x)$ il cui grafico è rappresentato nella figura 5. Il valore esatto è dato dall'area del trapezoide curvilineo $aABb$. Sia C il punto del grafico di ascissa $c = (a + b)/2$. Facciamo passare per A, B e C una parabola di secondo grado. Come sappiamo dal paragrafo precedente, essa è il grafico di un polinomio di secondo grado definito dall'equazione

$$P(x) = \frac{1}{2h^2}[(x-c)(x-b)y_0 - 2(x-a)(x-b)y_1 + (x-a)(x-c)y_2],$$

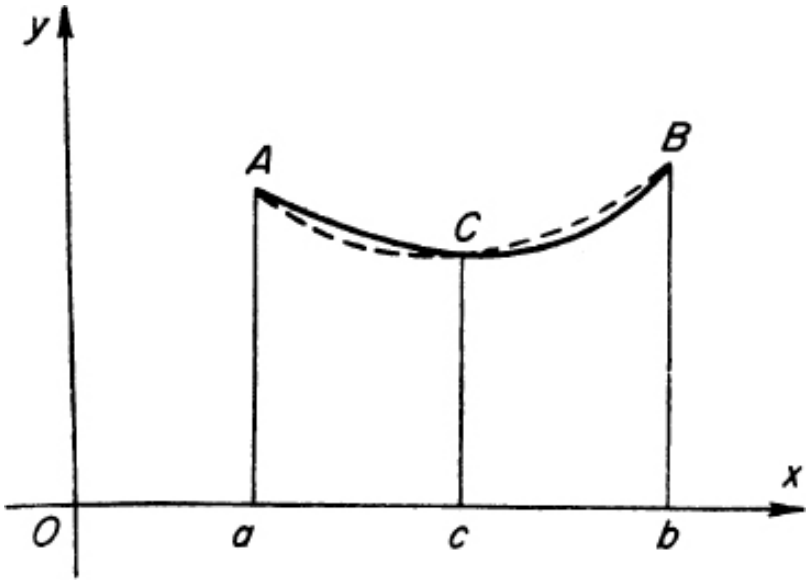


Figura 5.

dove

$$h = (b-a)/2,\ y_0 = f(a),\ y_1 = f(c),\ y_2 = f(b).$$

Nel linguaggio del paragrafo precedente, il polinomio $P(x)$ interpola $f(x)$ nei punti di ascissa a, c, b. Se il grafico di $f(x)$ non varia troppo bruscamente nell'intervallo $[a, b]$ e l'intervallo non è ampio, il polinomio $P(x)$ differirà poco dalla funzione in tutto l'intervallo. Anche i rispettivi integrali su $[a, b]$ saranno allora poco diversi. Possiamo dunque considerarli approssimativamente uguali:

$$\int_a^b f(x)\,dx \approx \int_a^b P(x)\,dx,$$

oppure, come si dice abitualmente, assumere il secondo integrale come approssimazione del primo. Semplici calcoli, che lasciamo al lettore, mostrano che

$$\int_a^b (x-c)(x-b)\,dx = \frac{2}{3}h^3, \quad -\int_a^b (x-a)(x-b)\,dx = \frac{4}{3}h^3,$$
$$\int_a^b (x-a)(x-c)\,dx = \frac{2}{3}h^3.$$

Pertanto

$$\int_a^b P(x)\,dx = (h/3)[f(a) + 4f(c) + f(b)].$$

L'integrale definito si può dunque calcolare mediante la seguente formula approssimata:

$$\int_a^b f(x)\,dx \approx \frac{h}{3}[f(a) + 4f(c) + f(b)].$$

Questa è la formula di Simpson.

Applichiamola, per esempio, all'integrale di sin x sull'intervallo [0, π]. In questo caso

$$h = \frac{\pi}{2}, \quad f(a) = \sin 0 = 0, \quad f(c) = \sin\frac{\pi}{2} = 1, \quad f(b) = \sin\pi = 0,$$

e di conseguenza $(h/3)[f(a) + 4f(c) + f(b)] = \frac{2}{3}\pi = 2.09\cdots$. D'altra parte l'integrale si può calcolare esattamente:

$$\int_0^\pi \sin x\,dx = -\cos x\Big|_0^\pi = 2.$$

L'errore non supera 0,1.

Se dividiamo [0, π] in due parti uguali e applichiamo separatamente la formula a ciascuna di esse, otteniamo

$$\int_0^{\pi/2} \sin x\,dx \approx (\pi/12)[\sin 0 + 4\sin(\pi/4) + \sin(\pi/2)]$$
$$= (\pi/12)(2\sqrt{2} + 1) \approx 1{,}00228;$$
$$\int_{\pi/2}^{\pi} \sin x\,dx \approx 1{,}00228.$$

In questo modo

$$\int_0^{\pi} \sin x\,dx \approx (\pi/6)(2\sqrt{2} + 1) \approx 2{,}00456.$$

e l'errore è ora circa 0,00456, quindi inferiore a 0,005.

In pratica, per calcolare approssimativamente l'integrale di $f(x)$ su $[a, b]$, dividiamo l'intervallo in un numero pari n di parti uguali mediante i punti $a = x_0 < x_1 < \cdots < x_n = b$. Applichiamo poi la formula di Simpson all'intervallo $[x_0, x_2]$, quindi a $[x_2, x_4]$ e così via. Si ottiene la formula composta di Simpson:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{3n}\,[f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \cdots + f(x_n)]. \tag{6}$$

Diamo ora, senza dimostrazione, la stima classica dell'errore. Se nell'intervallo $[a, b]$ la funzione $f(x)$ possiede una derivata quarta che soddisfa la disuguaglianza $|f^{IV}(x)| \leqslant M$, vale la stima

$$\left| \int_a^b f(x)\,dx - L(f) \right| \leqslant \frac{M(b-a)^5}{180n^4}. \tag{7}$$

Qui $L(f)$ indica il secondo membro della formula (6). L'errore è dunque di ordine n^{-4}.[3]

Avremmo potuto dividere $[a, b]$ in n parti uguali e assumere come approssimazione dell'integrale la somma delle aree dei rettangoli disegnati nella figura 6. Avremmo così ottenuto la formula dei rettangoli al punto medio:[4]

$$\int_a^b f(x)\,dx \approx \frac{b-a}{n}\,[f(x_0) + f(x_1) + \cdots + f(x_{n-1})]. \tag{8}$$

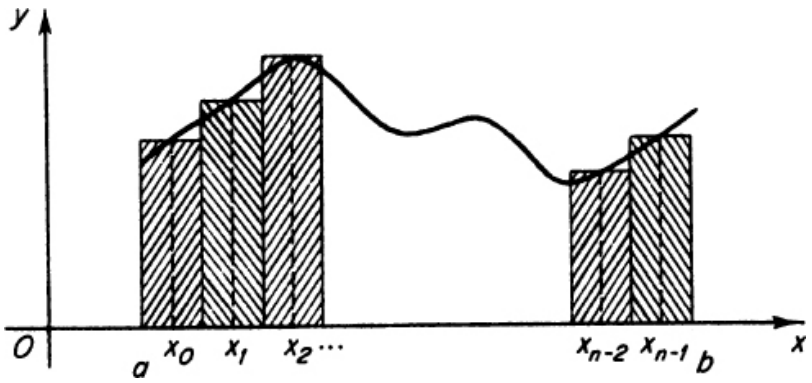


Figura 6.

Si può dimostrare che l'errore è di ordine n^{-2}, purché la funzione abbia derivata seconda limitata su $[a, b]$. Possiamo anche assumere come approssimazione la somma delle aree dei trapezi della figura 7, ottenendo la formula dei trapezi:

$$\int_a^b f(x)\,dx \approx \frac{b-a}{2n}\,[f(x_0) + 2f(x_1) + \cdots + 2f(x_{n-1}) + f(x_n)] \tag{9}$$

Anche in questo caso l'errore è di ordine n^{-2}, purché la derivata seconda sia limitata.

Si dice abitualmente che la formula di Simpson è più precisa delle formule dei trapezi e dei rettangoli. Senza una precisazione, tuttavia, questa affermazione non è vera. Se sappiamo

soltanto che la funzione ha derivata prima limitata, tutti e tre i metodi garantiscono un errore di ordine n^{-1}: sotto questa ipotesi la formula di Simpson non presenta un vantaggio essenziale. Per funzioni con derivata seconda limitata, sia la formula dei trapezi sia quella di Simpson garantiscono un errore di ordine n^{-2}. Se esistono derivate terza o quarta limitate, la stima generale per rettangoli e trapezi resta di ordine n^{-2}, mentre per Simpson diventa rispettivamente di ordine n^{-3} o n^{-4}. Quest'ultimo ordine è il migliore garantito in generale dalla formula di Simpson: la sola esistenza di derivate di ordine superiore non lo migliora. Se, per esempio, vogliamo sfruttare una derivata quinta limitata per ottenere un errore di ordine n^{-5}, occorre un diverso metodo di approssimazione dell'integrale. Per comprenderne la costruzione osserviamo quanto segue.

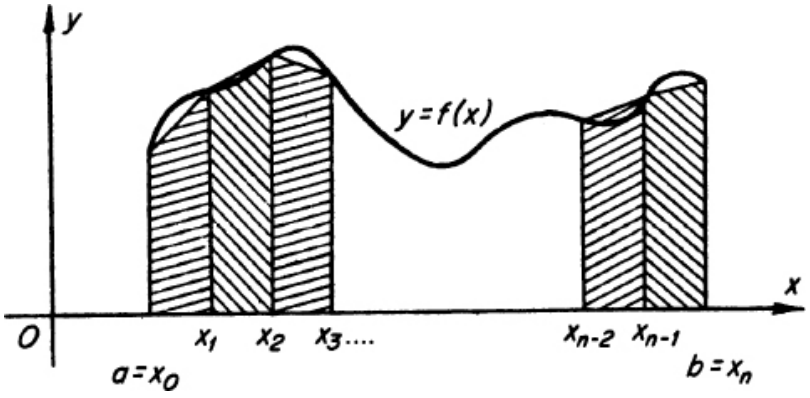


Figura 7.

Le formule dei trapezi e dei rettangoli sono esatte per i polinomi di primo grado: sostituendo alla funzione l'espressione $A + Bx$, con A e B costanti, si ottiene un'uguaglianza esatta. La formula di Simpson è invece esatta per tutti i polinomi di grado non superiore a tre, $A + Bx + Cx^2 + Dx^3$. È questo il punto essenziale. Supponiamo di dividere $[a, b]$ in n parti uguali e di applicare in ciascuna parte la stessa regola di quadratura, opportunamente riscalata, esatta per i polinomi $A + Bx + \cdots + Fx^{m-1}$. Per ogni funzione con derivata m-esima limitata, l'errore sarà di ordine n^{-m}. Se il grado di esattezza della regola è precisamente $m - 1$, la sola maggiore derivabilità della funzione non permette, in generale, di migliorare tale ordine garantito; particolari funzioni possono naturalmente dare errori minori o nulli.

Queste osservazioni mostrano l'importanza di trovare regole d'integrazione quanto più semplici possibile che siano esatte per i polinomi di un grado assegnato. La questione interessa da lungo tempo i matematici e possiede un'ampia letteratura. Qui possiamo soltanto richiamare alcuni risultati classici.

Sia assegnata la funzione $p(x)$. Come distribuire nell'intervallo $[-1, 1]$ i nodi $x_1, \ldots, x_m$ e come scegliere il numero K affinché l'uguaglianza

$$\int_{-1}^{1} f(x)\, p(x)\, dx = K \sum_{1}^{m} f(x_i),$$

valga per ogni polinomio $f(x)$ di grado non superiore a m?

Si può dimostrare che, per $p(x) = (1 - x^2)^{-\frac{1}{2}}$, il problema è risolto ponendo $K = \pi/m$ e scegliendo come x_i gli zeri del polinomio di Čebyšev $\cos(m \arccos x)$; si veda il §5.

Nel caso $p(x) = 1$, Čebyšev risolse il problema per $m = 1, 2, \ldots, 7$. Per $m = 8$ non esiste una soluzione con nodi reali ammissibili: la soluzione algebrica conduce a nodi complessi. Per $m = 9$ una soluzione esiste di nuovo. S. N. Bernštejn dimostrò però che, per ogni $m > 9$, il problema non ammette soluzione con tutti i nodi nell'intervallo $[-1, 1]$.

Una formula di quadratura esatta per i polinomi di grado non superiore a n si costruisce molto semplicemente mediante la formula di Lagrange (4). Integrandone i due membri su $[a, b]$, otteniamo

$$\int_a^b P_n(x)\,dx = \sum_{k=0}^{n} p_k f(x_k), \qquad (10)$$

dove

$$p_k = \int_a^b \frac{(x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_n)}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_n)}\,dx$$
$$(k = 0, 1, \cdots, n).$$

L'uguaglianza (10) vale per tutti i polinomi di grado non superiore a n; dunque la formula di quadratura

$$\int_a^b f(x)\,dx \approx \sum_0^n p_k f(x_k)$$

è esatta per tutti questi polinomi.

Quando

$$x_0 = a, \qquad x_1 = \frac{a+b}{2}, \qquad x_2 = b,$$

la formula si riduce, come abbiamo già visto, a quella di Simpson.

La distribuzione dei nodi d'interpolazione x_k, con $k = 0, 1, \ldots, n$, nell'intervallo $[a, b]$ può essere modificata. A ogni distribuzione corrisponde una formula di quadratura.

Gauss, il celebre matematico tedesco dell'Ottocento, dimostrò che è possibile distribuire gli $n + 1$ nodi x_k in modo che la formula sia esatta per tutti i polinomi di grado non soltanto n, ma fino a $2n + 1$.

Il polinomio

$$A_{n+1}(x) = (x - x_0)(x - x_1)\cdots(x - x_n)$$

di grado $n + 1$, costruito con i nodi di Gauss x_k, possiede una proprietà notevole: per ogni polinomio $P(x)$ di grado inferiore a $n + 1$ vale l'uguaglianza

$$\int_a^b A_{n+1}(x)\, P(x)\, dx = 0.$$

In altri termini, $A_{n+1}(x)$ è ortogonale, nell'intervallo $[a, b]$, a tutti i polinomi di grado non superiore a n. Questi sono i polinomi di Legendre relativi all'intervallo $[a, b]$, qui normalizzati in modo che il coefficiente del termine di grado massimo sia uno.

§4. Il concetto di migliore approssimazione uniforme di Čebyšev

Impostazione del problema.

Čebyšev giunse all'idea di migliore approssimazione uniforme partendo da un problema strettamente pratico. Egli fu infatti non soltanto uno dei maggiori matematici dell'Ottocento, fondatore di varie discipline poi ampiamente sviluppate, ma anche un eminente ingegnere del suo tempo. Si interessò in particolare alla costruzione di meccanismi capaci di produrre una traiettoria di movimento assegnata. Spieghiamo questa idea.

Sia data la curva $y = f(x)$ sull'intervallo $a \leqslant x \leqslant b$. Si vuole costruire un meccanismo che soddisfi determinati requisiti tecnici e nel quale un certo punto descriva, durante il funzionamento, questa curva con la massima precisione possibile. Čebyšev procedeva nel modo seguente. Dapprima, affrontando la questione da ingegnere, ideava un meccanismo che desse una prima approssimazione della traiettoria richiesta. Un certo punto A del meccanismo, ancora provvisorio, descriveva così una curva

$$y = \phi(x), \qquad (11)$$

simile soltanto nelle sue linee generali alla curva desiderata $y = f(x)$. Il meccanismo è composto di parti distinte: ingranaggi, leve di vario tipo e così via. Tutte queste parti hanno dimensioni determinate,

$$\alpha_0 , \alpha_1 , \alpha_2 , \cdots , \alpha_m , \qquad (12)$$

che descrivono interamente il meccanismo e, di conseguenza, la curva (11). Esse costituiscono i parametri del meccanismo e della curva.[5] La curva (11) dipende dunque non soltanto dalla variabile x, ma anche dai parametri (12). A ogni sistema assegnato di valori dei parametri corrisponde una curva determinata, la cui equazione si può scrivere convenientemente nella forma

$$y = \phi(x; \alpha_0 , \alpha_1 , \cdots , \alpha_m). \qquad (13)$$

In questi casi si dice che abbiamo ottenuto una famiglia di funzioni (13), definite sull'intervallo $a \leqslant x \leqslant b$ e dipendenti dagli $m + 1$ parametri (12).

Per proseguire nella soluzione, Čebyšev operava da matematico puro. In modo del tutto naturale, propose di misurare lo scostamento della funzione $f(x)$ dalla funzione approssimante $\phi(x; \alpha_0, \alpha_1, \ldots, \alpha_m)$ mediante la quantità

$$\|f - \phi\| = \max_{a \leqslant x \leqslant b} |f(x) - \phi(x;\ \alpha_0, \alpha_1, \cdots, \alpha_m)|, \qquad (14)$$

uguale al massimo del valore assoluto della differenza $f(x) - \phi(x; \alpha_0, \alpha_1, \ldots, \alpha_m)$ sull'intervallo $a \leqslant x \leqslant b$ (figura 8). Questa quantità è evidentemente una funzione

$$\|f - \phi\| = F(\alpha_0, \alpha_1, \cdots, \alpha_m) \qquad (15)$$

dei parametri $\alpha_0, \alpha_1, \ldots, \alpha_m$. Occorre ora trovare i valori dei parametri che rendono minima la funzione (15). Tali valori determinano una funzione ϕ detta migliore approssimante uniforme di $y = f(x)$ fra tutte le funzioni della famiglia considerata. Il valore assunto da $F(\alpha_0, \alpha_1, \ldots, \alpha_m)$ per questi parametri si chiama errore di migliore approssimazione uniforme di $f(x)$ su $[a, b]$ mediante le funzioni della famiglia (13), o, più brevemente, migliore approssimazione uniforme. Di solito lo si indica con $E_m(f)$. Il termine «uniforme» viene spesso sostituito, soprattutto nella letteratura non sovietica dell'epoca, dall'espressione «di Čebyšev». Entrambe le denominazioni sottolineano il carattere particolare di questa approssimazione: sono infatti possibili altri criteri, per esempio quello della migliore approssimazione in media quadratica mediante funzioni di una famiglia assegnata. Ne parleremo nel §8.

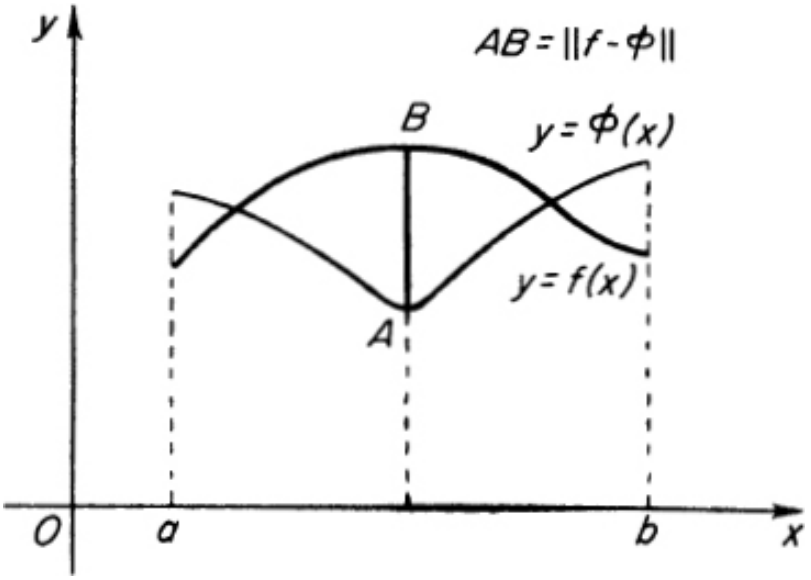


Figura 8.

Čebyšev scoprì per primo varie leggi che governano il tipo di approssimazione qui considerato. Trovò che, in molti casi, la funzione ϕ che meglio approssima uniformemente $f(x)$ su $[a, b]$ possiede la seguente proprietà notevole: il massimo (15) del valore assoluto della differenza

$$f(x) - \phi(x; \alpha_0, \alpha_1, \cdots, \alpha_m)$$

è raggiunto in almeno $m + 2$ punti ordinati dell'intervallo $[a, b]$, nei quali la differenza assume segni alterni (figura 9).

Non possiamo formulare qui con precisione le condizioni che rendono valida questa proposizione. Rimandiamo i lettori più preparati all'articolo di V. L. Gončarov «La teoria della migliore approssimazione delle funzioni», nel primo volume dell'«Eredità scientifica di Čebyšev».

Il caso dell'approssimazione mediante polinomi.

Le ricerche di Čebyšev acquistano particolare importanza per la teoria generale dell'approssimazione quando si applicano a una funzione continua arbitraria $f(x)$ su $[a, b]$ e ai polinomi $P_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ di grado non superiore a n. Questi polinomi costituiscono una famiglia dipendente da $n + 1$ parametri, i coefficienti. La teoria di Čebyšev è pienamente applicabile a tale famiglia. Per verificare se un dato polinomio $P_n(x)$ fornisca la migliore approssimazione uniforme, cerchiamo dunque i punti in cui $|f(x) - P_n(x)|$ raggiunge il suo massimo L. Se $L > 0$, il criterio richiede che fra questi punti se ne possano trovare $n + 2$, ordinati da sinistra a destra, $x_1 < x_2 < \cdots < x_{n+2}$, nei quali la differenza cambi successivamente segno:

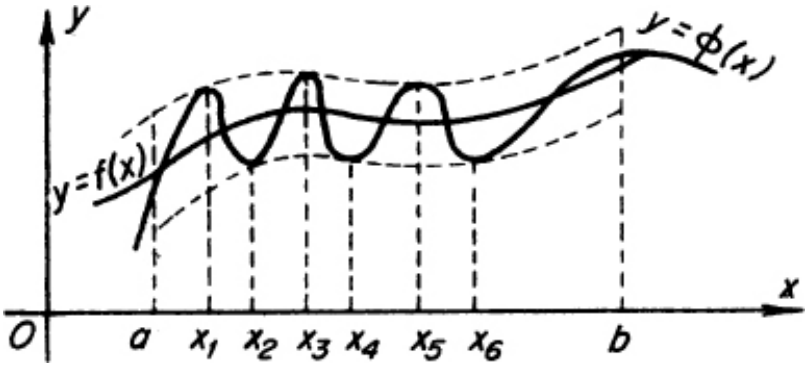


Figura 9.

$$f(x_1) - P_n(x_1) = \pm L,$$
$$f(x_2) - P_n(x_2) = \pm L,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$f(x_{n+2}) - P_n(x_{n+2}) = \pm (-1)^{n+1} L,$$

In questo caso $P_n(x)$ è il polinomio migliore, e altrimenti non lo è. Il caso $L = 0$ corrisponde naturalmente a un'approssimazione già esatta. Per esempio, la migliore approssimazione uniforme della funzione rappresentata nella figura 10 mediante polinomi di primo grado $P_1(x) = p + qx$ è il polinomio $p_0 + q_0x$ il cui grafico è parallelo alla corda AB e divide in due parti uguali il parallelogramma compreso fra la corda e la tangente parallela CD. Infatti, il valore assoluto della differenza $f(x) - (p_0 + q_0x)$ assume il massimo nei punti di ascissa $x_0 = a$, x_1 e $x_2 = b$, dove x_1 è l'ascissa del punto di tangenza F; in questi punti la differenza cambia successivamente segno. Per evitare equivoci, consideriamo qui una curva con la convessità

mostrata in figura e dotata di tangente in ogni punto. In questo esempio $E_1(f)$ è la metà della lunghezza di uno qualsiasi dei segmenti uguali *AC*, *BD* e *GF*.

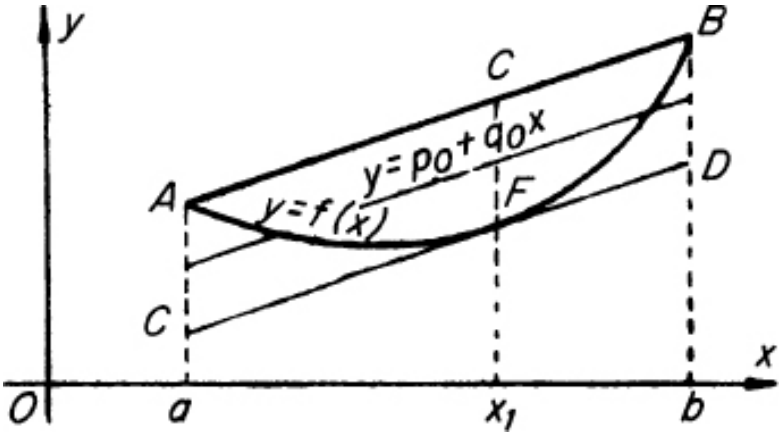


Figura 10.

§5. I polinomi di Čebyšev che si discostano meno da zero

Consideriamo il seguente problema: trovare il polinomio $P_{n-1}(x)$, di grado non superiore a $n - 1$, che meglio approssima uniformemente la funzione x^n nell'intervallo $[-1, 1]$.

Il polinomio cercato soddisfa l'uguaglianza

$$x^n - P_{n-1}(x) = \frac{1}{2^{n-1}} \cos n \arccos x. \qquad (16)$$

Questo fatto discende direttamente dal teorema di Čebyšev, purché dimostriamo tre proprietà del secondo membro della (16): esso è un polinomio algebrico di grado n con coefficiente di x^n uguale a uno; il suo valore assoluto raggiunge, nell'intervallo $[-1, 1]$, il massimo $L = 1/2^{n-1}$ negli $n + 1$ punti $x_k = \cos(k\pi/n)$, con $k = 0, 1, \ldots, n$; infine, i valori assunti in questi punti hanno segni alterni.

Per dimostrare che il secondo membro della (16) è un polinomio di grado n con coefficiente principale uno, procediamo così.

Supponiamo di avere già dimostrato, per un dato numero naturale n, che

$$\cos n \arccos x = 2^{n-1}[x^n - Q_{n-1}(x)];$$
$$-\sqrt{1 - x^2} \sin n \arccos x = 2^{n-1}[x^{n+1} - Q_n(x)],$$

dove Q_{n-1} e Q_n sono polinomi algebrici di grado rispettivamente non superiore a $n - 1$ e a n. Le analoghe uguaglianze valgono allora anche al passo successivo, come si ricava facilmente dalle formule

$$\cos (n+1) \arccos x = x \cos n \arccos x - \sqrt{1-x^2} \sin n \arccos x;$$
$$-\sqrt{1-x^2} \sin (n+1) \arccos x$$
$$= -x\sqrt{1-x^2} \sin n \arccos x + (x^2-1) \cos n \arccos x,$$

Ma le uguaglianze sono vere per $n = 1$, poiché

$$\cos \arccos x = x,$$
$$-\sqrt{1-x^2} \sin \arccos x = x^2 - 1.$$

Sono dunque vere per ogni n.

Il secondo membro della (16) si chiama polinomio di Čebyšev di grado n che si discosta meno da zero: fra tutti i polinomi monici dello stesso grado, cioè con coefficiente principale uno, è quello di minima norma uniforme su $[-1, 1]$. Čebyšev fu il primo a porre e risolvere il problema. I primi di questi polinomi sono

$$T_0(x) = 1,$$
$$T_1(x) = x,$$
$$T_2(x) = \tfrac{1}{2}(2x^2 - 1),$$
$$T_3(x) = \tfrac{1}{4}(4x^3 - 3x),$$
$$T_4(x) = \tfrac{1}{8}(8x^4 - 8x^2 + 1),$$
$$T_5(x) = \tfrac{1}{16}(16x^5 - 20x^3 + 5x).$$

Abbiamo già visto il ruolo importante dei polinomi di Čebyšev nell'interpolazione e nell'integrazione approssimata. Aggiungiamo ancora alcune osservazioni sull'interpolazione.

Se l'errore della migliore approssimazione non è nullo, la differenza fra una funzione continua $f(x)$ e il suo migliore polinomio approssimante $P_n(x)$ assume segni alterni in $n + 2$ punti. Per la continuità, fra questi punti si trovano almeno $n + 1$ punti nei quali la differenza si annulla. Il polinomio migliore è quindi anche un polinomio interpolatore, con una particolare scelta dei nodi. Se l'errore è nullo, la coincidenza vale in tutto l'intervallo.

Il problema della migliore approssimazione uniforme di una funzione continua può dunque essere ricondotto alla scelta, su $[-1, 1]$, di un sistema di nodi $x_0, x_1, \ldots, x_n$ tale che il corrispondente polinomio interpolatore Q renda minimo lo scostamento $||f - Q|| = \max_x |f(x) - Q(x)|$. Purtroppo questi nodi sono spesso difficili da trovare in pratica. Occorre di solito risolvere il problema in modo approssimato, e qui i polinomi di Čebyšev svolgono un ruolo speciale. Se scegliamo come nodi gli zeri di $\cos[(n + 1) \arccos x]$, il polinomio interpolatore ottenuto fornisce, almeno per grandi n e per funzioni sufficientemente regolari, uno scostamento uniforme che differisce poco dall'errore della migliore approssimazione polinomiale. In numerosi casi importanti, l'espressione alquanto vaga «differisce poco» può essere sostituita da stime quantitative precise, che non ricaveremo qui.

Torniamo al polinomio di Čebyšev, ora con una diversa normalizzazione: $T_n(x) = M \cos(n \arccos x)$, $(-1 \leqslant x \leqslant 1)$, dove M è un numero positivo. Il suo valore assoluto non supera evidentemente M nell'intervallo $[-1, 1]$. La sua derivata è

$$T'_n(x) = nM \sin(n \arccos x)/\sqrt{(1 - x^2)}, \quad -1 < x < 1,$$

e soddisfa, per $-1 < x < 1$, la disuguaglianza

$$| T'_n(x)| \leqslant \frac{nM}{\sqrt{1 - x^2}} .$$

Questa disuguaglianza vale in realtà per ogni polinomio $P_n(x)$ di grado non superiore a n il cui valore assoluto non superi M su $[-1, 1]$. Per la derivata di un tale polinomio, nei punti interni $-1 < x < 1$, si ha quindi

$$| P'_n(x)| \leqslant \frac{nM}{\sqrt{1 - x^2}} .$$

La disuguaglianza è dovuta ad A. A. Markov: discende direttamente da suoi risultati ancora più ampi. Markov la ottenne studiando una questione suggeritagli da D. I. Mendeleev.

Nel 1912 S. N. Bernštejn ottenne per i polinomi trigonometrici una disuguaglianza analoga, che porta il suo nome. Servendosi di queste disuguaglianze, mostrò per primo come dedurre proprietà di derivabilità di una funzione dalla rapidità con cui tende a essa la successione delle sue migliori approssimazioni. Alcuni risultati di questo tipo sono esposti nei §§6 e 7.

§6. Il teorema di Weierstrass e il legame fra migliore approssimazione e derivabilità

Il teorema di Weierstrass.

Applicando ai polinomi la definizione generale data nel §4, si ottiene la seguente definizione. L'errore di migliore approssimazione uniforme di $f(x)$ su $[a, b]$ mediante polinomi di grado non superiore a n è il numero non negativo $E_n(f)$ uguale al minimo dell'espressione

$$\max_{a \leqslant x \leqslant b} |f(x) - P_n(x)| = \| f - P_n \| ,$$

al variare di $P_n(x)$ fra tutti i polinomi di grado non superiore a n.

Anche quando non sappiamo determinare esattamente il polinomio migliore, è di grande interesse pratico e teorico stimare quanto più precisamente possibile $E_n(f)$. Se vogliamo approssimare f con precisione δ, ossia in modo che

$$|f(x) - P_n(x)| \leqslant \delta \qquad (17)$$

per ogni x dell'intervallo, non ha senso cercare fra i polinomi di grado non superiore a n quando $E_n(f) > \delta$: nessuno di essi soddisferà la (17). Se invece sappiamo che $E_n(f) < \delta$, la ricerca ha senso, poiché polinomi con la precisione richiesta esistono certamente.

Le proprietà delle migliori approssimazioni per diverse classi di funzioni sono state studiate in modo approfondito. Ricordiamo anzitutto il seguente fatto fondamentale.

Se una funzione f(x) è continua sull'intervallo chiuso e limitato [a, b], la sua migliore approssimazione $E_n(f)$ tende a zero quando n tende all'infinito.

È il teorema dimostrato da Weierstrass alla fine dell'Ottocento. Esso garantisce la possibilità di approssimare una funzione continua arbitraria mediante un polinomio con qualsiasi precisione desiderata. L'insieme di tutti i polinomi sta dunque all'insieme delle funzioni continue sull'intervallo in una relazione analoga a quella fra l'insieme R dei numeri razionali e l'insieme H di tutti i numeri reali. Infatti, per ogni irrazionale α e ogni numero positivo ε, piccolo quanto si vuole, esiste un razionale r tale che $|\alpha - r| < \varepsilon$. Analogamente, se $f(x)$ è continua su $[a, b]$ e ε è un numero positivo arbitrario, il teorema di Weierstrass garantisce un polinomio $P_n(x)$ per il quale $|f(x) - P_n(x)| < \varepsilon$ in tutto l'intervallo. Quindi $E_n(f)$ tende a zero per $n \to \infty$.

Illustriamo geometricamente il teorema. Dato il grafico di una funzione continua su $[a, b]$ e un numero positivo ε, piccolo quanto si vuole, circondiamo il grafico con una striscia di spessore verticale 2ε, in modo che esso passi al centro della striscia. È sempre possibile scegliere un polinomio algebrico

$$P_n(x) = a_0 + a_1 x + \cdots + a_n x^n,$$

di grado abbastanza elevato il cui grafico sia interamente contenuto nella striscia.

Osserviamo ancora quanto segue. Sia $f(x)$ una funzione continua su $[a, b]$ e siano $P_n(x)$, con $n = 1, 2, \ldots$, i suoi migliori polinomi approssimanti uniformi. Si vede facilmente che $f(x)$ si può rappresentare mediante la serie $f(x) = P_1(x) + [P_2(x) - P_1(x)] + [P_3(x) - P_2(x)] + \cdots$, uniformemente convergente su $[a, b]$. Infatti, la somma dei primi n termini è $P_n(x)$ e

$$\max_{a \leqslant x \leqslant b} |f(x) - P_n(x)| = E_n(f),$$

mentre $E_n(f) \to 0$ per $n \to \infty$.

Otteniamo così una nuova formulazione del teorema di Weierstrass:

Ogni funzione continua sull'intervallo [a, b] può essere rappresentata mediante una serie di polinomi algebrici uniformemente convergente alla funzione.

Questo risultato ha grande importanza teorica. Garantisce la possibilità di rappresentare una funzione continua arbitraria, comunque assegnata - per esempio mediante un grafico - con un'espressione analitica. Qui «espressione analitica» è inteso nel senso storico di funzione elementare oppure di funzione ottenuta da una successione di funzioni elementari mediante un passaggio al limite: non significa necessariamente «funzione analitica» nel senso moderno.

Il risultato mise definitivamente in crisi la concezione rimasta diffusa quasi fino alla metà dell'Ottocento. Diciamo «definitivamente» perché il teorema di Weierstrass era stato preceduto da risultati generali simili, soprattutto sulle serie di Fourier. Prima di questi sviluppi si riteneva che le espressioni analitiche rappresentassero funzioni dotate delle proprietà particolarmente favorevoli proprie delle funzioni analitiche; si dava di solito per scontato che fossero derivabili infinite volte e persino sviluppabili in serie di potenze. Queste convinzioni si rivelarono prive di fondamento. Una funzione può non avere derivata in alcun punto del suo intervallo di definizione ed essere tuttavia rappresentabile con un'espressione analitica nel senso appena precisato.

Dal punto di vista metodologico, il valore della scoperta sta nel mostrare con piena chiarezza che, almeno in linea di principio, i metodi matematici si applicano a una classe di leggi immensamente più vasta di quanto si fosse immaginato.

Sono note numerose dimostrazioni del teorema di Weierstrass. Per lo più esse costruiscono, a partire da una funzione continua f, una successione di polinomi che l'approssimano uniformemente al crescere del grado. Il polinomio di semplice costruzione

$$B_n(x) = \sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} f(k/n),$$

approssima una funzione continua $f(x)$ sull'intervallo [0, 1] e si chiama polinomio di Bernštejn. Al crescere di n, converge uniformemente su [0, 1] alla funzione da cui è costruito.[6] Qui C_n^k è il numero delle combinazioni di n elementi presi k alla volta.

Un teorema analogo a quello di Weierstrass vale anche nel dominio complesso. Risultati di grande ampiezza in questa direzione si devono a M. A. Lavrent'ev, M. V. Keldyš e S. N. Mergeljan.

Legame fra ordine di approssimazione e derivabilità.

Ricordiamo i risultati seguenti. Se una funzione $f(x)$ possiede su $[a, b]$ una derivata $f^{(r)}(x)$ di ordine r il cui valore assoluto non supera K, allora la migliore approssimazione soddisfa la disuguaglianza

$$E_n(f) \leqslant \frac{c_r K}{n^r}, \qquad (18)$$

dove, fissato l'intervallo, c_r è una costante dipendente soltanto da r (teorema di Jackson); se si varia anche l'intervallo, occorre tener conto della sua lunghezza. La (18) mostra che l'errore $E_n(f)$ tende a zero tanto più rapidamente quanto più alto è l'ordine delle derivate disponibili. In altri termini, quanto più la funzione è regolare, tanto più rapida è la decrescita della migliore approssimazione. Bernštejn dimostrò che, in un senso opportuno, vale anche il viceversa.

Ancora più favorevole è il caso delle funzioni analitiche in un intorno dell'intervallo chiuso considerato. Bernštejn dimostrò che per esse vale la stima

$$E_n(f) \leqslant cq^n, \qquad (19)$$

dove c e q dipendono dalla funzione f e $0 < q < 1$: l'errore è dunque maggiorato dai termini di una progressione geometrica decrescente. Dimostrò inoltre il viceversa: la disuguaglianza (19) implica l'analiticità di f sull'intervallo.

Abbiamo enunciato alcuni risultati fondamentali, scoperti all'inizio del Novecento, che hanno orientato la ricerca nella teoria dell'approssimazione. Il loro valore pratico risulta dall'esempio seguente.

Sia $Q_n(x)$ il polinomio di grado non superiore a n che interpola $f(x)$ su $[-1, 1]$ negli $n + 1$ zeri di $\cos[(n + 1) \arccos x]$. In tutto l'intervallo vale la stima $|f(x) - Q_n(x)| \leq c \ln(n + 1) E_n(f)$, per $n \geq 1$, con c indipendente da n. Se f è abbastanza regolare, possiamo sostituire a $E_n(f)$ le maggiorazioni (18) o (19), ottenendo una buona stima dell'errore d'interpolazione. Poiché il logaritmo cresce molto lentamente, l'ordine della stima differisce poco dall'ordine di decrescita della migliore approssimazione. Il vantaggio dei nodi di Čebyšev è che, con altre scelte dei nodi, il fattore logaritmico può essere sostituito da un fattore che cresce molto più rapidamente; ciò avviene in particolare per i nodi equidistanti.

§7. Serie di Fourier

Origine delle serie di Fourier.

Le serie di Fourier nacquero dallo studio di alcuni fenomeni fisici, in particolare delle piccole oscillazioni dei mezzi elastici. Un esempio caratteristico è quello di una corda musicale. Lo studio delle corde vibranti fu precisamente l'origine storica delle serie di Fourier e determinò la direzione di sviluppo della loro teoria.

Consideriamo una corda tesa, fissata agli estremi nei punti $x = 0$ e $x = l$ dell'asse Ox (figura 11). Se la spostiamo dalla posizione di equilibrio, essa si mette a oscillare.

Seguiamo il moto di un punto determinato della corda, di ascissa x_0. Il suo spostamento verticale dalla posizione di equilibrio è una funzione $\phi(t)$ del tempo. Si può dimostrare che è possibile assegnare alla corda, all'istante $t = 0$, posizione e velocità iniziali tali che il punto considerato compia oscillazioni armoniche verticali descritte dalla funzione

$$\phi = \phi(t) = A \cos \alpha k t + B \sin \alpha k t. \qquad (20)$$

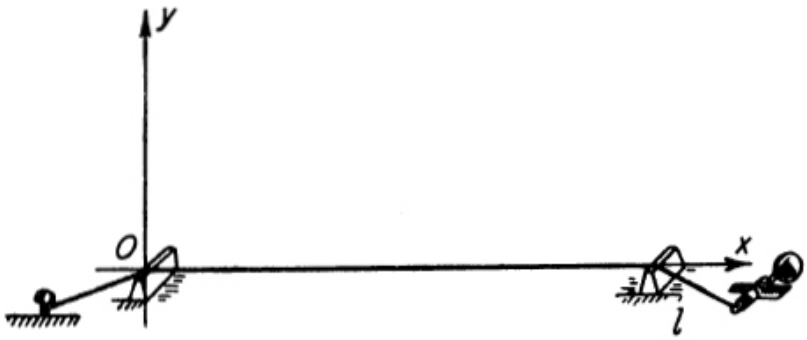


Figura 11.

Qui α è una costante che dipende soltanto dalle caratteristiche fisiche della corda - densità, tensione e lunghezza -, k è un intero positivo e A, B sono costanti.

Il ragionamento riguarda soltanto piccole oscillazioni. Possiamo perciò assumere approssimativamente che ogni punto x_0 si muova solo in direzione verticale, trascurandone gli spostamenti orizzontali.[7] Supponiamo inoltre trascurabile l'attrito che accompagna il movimento. Con queste ipotesi approssimate, le oscillazioni non si smorzano.

I moti periodici descritti dalle funzioni armoniche (20) non esauriscono naturalmente le possibili oscillazioni del punto x_0, ma possiedono una proprietà notevole. Gli esperimenti e la teoria che li interpreta mostrano che ogni oscillazione ammissibile si può ottenere sovrapponendo opportune oscillazioni armoniche della forma (20). Le oscillazioni relativamente semplici risultano dalla combinazione di un numero finito di tali moti e sono quindi descritte da funzioni della forma

$$\phi(t) = A_0 + \sum_{k=1}^{n} (A_k \cos \alpha k t + B_k \sin \alpha k t),$$

dove A_k e B_k sono opportune costanti. Queste funzioni si chiamano polinomi trigonometrici. Nei casi più complessi, l'oscillazione risulta dalla sovrapposizione di infinite oscillazioni della forma (20), corrispondenti a $k = 1, 2, 3, \ldots$, con costanti A_k e B_k opportunamente scelte. Siamo così condotti a rappresentare la funzione $\phi(t)$, di periodo $2\pi/\alpha$, che descrive il moto del punto, mediante una serie

$$\phi(t) = A_0 + \sum_{k=1}^{\infty} (A_k \cos \alpha k t + B_k \sin \alpha k t). \qquad (21)$$

In fisica si presentano molte altre situazioni nelle quali è naturale considerare una funzione assegnata, anche se non descrive direttamente un'oscillazione, come somma di una serie trigonometrica infinita della forma (21). Un esempio nasce proprio nello studio della corda vibrante. La legge delle oscillazioni successive, quando all'inizio dell'esperimento si è impressa alla corda una determinata deformazione - per esempio quella della figura 12 -, si calcola facilmente se si conosce lo sviluppo $f(x) = \sum_1^\infty a_k \sin (k\pi/l)x$, caso particolare della (21), della funzione $f(x)$ che descrive la posizione iniziale.

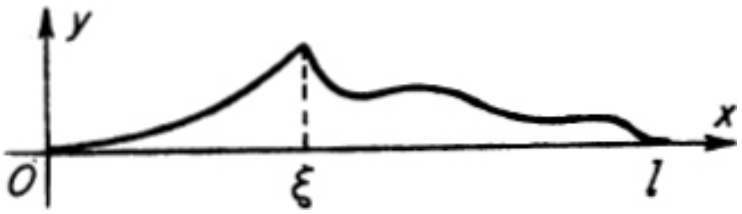


Figura 12.

Sviluppo delle funzioni in serie trigonometriche.

Sorge così la domanda fondamentale: quali funzioni di periodo $2\pi/\alpha$ possono essere rappresentate come somma di una serie della forma (21)? La questione fu posta nel Settecento da Eulero e Bernoulli, in relazione agli studi di quest'ultimo sulla corda vibrante. Guidato da considerazioni fisiche, Bernoulli sosteneva che una classe molto vasta di funzioni continue, comprendente in particolare tutti i grafici tracciabili a mano, potesse essere sviluppata in serie trigonometriche. Molti contemporanei criticarono aspramente questa opinione. Restavano legati all'idea che una funzione rappresentata da un'espressione analitica, quale una serie trigonometrica, dovesse avere buone proprietà di derivabilità. Ma la funzione della figura 12 non possiede nemmeno la derivata nel punto ξ: come potrebbe dunque essere definita da una sola espressione analitica su tutto [0, *l*]?

Sappiamo ormai che l'intuizione fisica di Bernoulli era corretta. Per chiudere la controversia occorse però un secolo intero: una risposta soddisfacente richiedeva anzitutto una definizione rigorosa di limite e di somma di una serie.

Le indagini matematiche fondamentali che confermarono il punto di vista fisico, pur basandosi ancora sulle precedenti concezioni dei fondamenti dell'analisi, furono compiute dal matematico francese Fourier negli anni 1807-1822.

Infine, nel 1829, il matematico tedesco Dirichlet dimostrò con pieno rigore che ogni funzione continua di periodo $2\pi/\alpha$,[8] con un numero finito di massimi e minimi in un periodo, ammette un unico sviluppo in serie trigonometrica di Fourier, uniformemente convergente[9] alla funzione.

La figura 13 rappresenta una funzione che soddisfa le condizioni di Dirichlet: il grafico è continuo e periodico, con periodo 2π, e presenta un massimo e un minimo nell'intervallo $0 \leqslant x \leqslant 2\pi$.

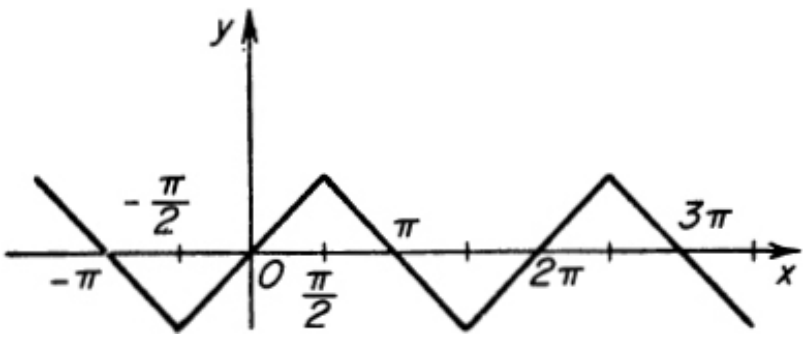


Figura 13.

I coefficienti di Fourier.

Per semplificare le formule, d'ora in poi considereremo funzioni di periodo 2π. Sia $f(x)$ una di queste funzioni, continua e soggetta alla condizione di Dirichlet. Il teorema di Dirichlet ne assicura lo sviluppo in serie trigonometrica

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \qquad (22)$$

uniformemente convergente alla funzione. Scrivere il primo termine come $a_0/2$ anziché come a_0 è soltanto una scelta di comodità, di cui vedremo presto la ragione.

Poniamoci il problema di calcolare i coefficienti a_k e b_k della serie a partire dalla funzione $f(x)$.

A questo scopo osserviamo le uguaglianze seguenti:

$$\begin{aligned}
&\int_{-\pi}^{\pi} \cos kx \cos lx\, dx = 0 \quad (k \neq l;\ k, l = 0, 1, \cdots),\\
&\int_{-\pi}^{\pi} \sin kx \sin lx\, dx = 0 \quad (k \neq l;\ k, l = 0, 1, \cdots),\\
&\int_{-\pi}^{\pi} \sin kx \cos lx\, dx = 0 \quad (k, l = 0, 1, 2, \cdots), \qquad (23)\\
&\int_{-\pi}^{\pi} \cos^2 kx\, dx = \pi \qquad (k = 1, 2, \cdots),\\
&\int_{-\pi}^{\pi} \sin^2 kx\, dx = \pi \qquad (k = 1, 2, \cdots),
\end{aligned}$$

Il lettore può verificarle facilmente trasformando i prodotti di funzioni trigonometriche in somme e differenze e i loro quadrati in espressioni contenenti funzioni dell'angolo doppio. La prima uguaglianza afferma che l'integrale, su un periodo, del prodotto di due funzioni distinte della successione 1, cos x, sin x, cos $2x$, sin $2x$, … è nullo: è la proprietà di ortogonalità delle funzioni trigonometriche. L'integrale del quadrato di ciascuna funzione non costante della successione è invece π. Fa eccezione la funzione identicamente uguale a uno, il cui quadrato ha integrale 2π sul periodo. Proprio per questo conviene scrivere il primo termine della (22) come $a_0/2$.

Ora il problema si risolve facilmente. Per calcolare a_m, moltiplichiamo il primo membro e ogni termine del secondo membro della (22) per cos mx, quindi integriamo termine a termine su un periodo 2π. Questo passaggio è lecito perché la serie resta uniformemente convergente dopo la moltiplicazione. Per la (23), tutti gli integrali al secondo membro sono nulli tranne quello relativo al termine in cos mx; pertanto

$$\int_{-\pi}^{\pi} f(x) \cos mx\, dx = a_m \pi,$$

da cui

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos mx\,dx \qquad (m = 0, 1, 2, \cdots). \qquad (24)$$

Analogamente, moltiplicando i due membri della (22) per sin *mx* e integrando sul periodo, otteniamo

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin mx\,dx \qquad (m = 1, 2, \cdots), \qquad (25)$$

e il problema è risolto. I numeri a_m e b_m, calcolati con le formule (24) e (25), si chiamano coefficienti di Fourier della funzione $f(x)$.

Consideriamo per esempio la funzione di periodo 2π della figura 13. Essa è continua e soddisfa la condizione di Dirichlet, quindi la sua serie di Fourier converge uniformemente alla funzione.

È facile vedere che $f(-x) = -f(x)$. La stessa proprietà vale per $F_1(x) = f(x)\cos mx$, il cui grafico è dunque simmetrico rispetto all'origine. Per ragioni geometriche $\int_{-\pi}^{\pi} F_1(x)\,dx = 0$, perciò $a_m = 0$ per $m = 0, 1, 2, \ldots$. La funzione $F_2(x) = f(x)\sin mx$ ha invece grafico simmetrico rispetto all'asse *Oy*, quindi

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} F_2(x)\,dx = \frac{2}{\pi}\int_0^{\pi} F_2(x)\,dx.$$

Per *m* pari, il grafico di F_2 nell'intervallo $[0, \pi]$ è simmetrico rispetto al punto $(\pi/2, 0)$: i valori in punti equidistanti dal centro hanno segno opposto. Ne segue $b_m = 0$. Per *m* dispari, $m = 2l + 1$ con $l = 0, 1, 2, \ldots$, il grafico è invece simmetrico rispetto alla retta $x = \pi/2$, e dunque

$$b_{2l+1} = \frac{4}{\pi}\int_0^{\pi/2} F_2(x)\,dx.$$

Come si vede dalla figura, su $[0, \pi/2]$ vale semplicemente $f(x) = x$. Integrando per parti si ottiene

$$b_{2l+1} = \frac{4}{\pi}\int_0^{\pi/2} x\sin(2l+1)x\,dx = \frac{4(-1)^l}{\pi(2l+1)^2},$$

e quindi

$$f(x) = (4/\pi)\sum_{l=0}^{\infty}[(-1)^l \sin((2l+1)x)/(2l+1)^2].$$

Abbiamo così trovato lo sviluppo della funzione in serie di Fourier.

Convergenza delle somme parziali di Fourier.

Nelle applicazioni si assume abitualmente come approssimazione della funzione $f(x)$, di periodo 2π, la somma parziale

$$S_n = \frac{a_0}{2} + \sum_1^n (a_k \cos kx + b_k \sin kx)$$

che comprende le armoniche fino all'ordine n. Sorge allora il problema di stimare l'errore. Se $f(x)$ possiede una derivata $f^{(r)}(x)$ di ordine r che soddisfa per ogni x la disuguaglianza $|f^{(r)}(x)| \leqslant K$, per $n \geq 2$ l'errore può essere stimato così:

$$|f(x) - S_n(x)| \leqslant \frac{c_r K \ln n}{n^r},$$

dove c_r dipende soltanto da r. L'errore tende a zero al crescere di n, tanto più rapidamente quanto più numerose sono le derivate della funzione.

Per una funzione periodica analitica sull'intero asse reale vale una stima ancora migliore:

$$|f(x) - S_n(x)| < cq^n, \qquad (26)$$

dove c e q sono costanti positive dipendenti da f, con $q < 1$. Vale anche il viceversa: se la (26) è soddisfatta, la funzione è necessariamente analitica. Questo risultato, scoperto all'inizio del Novecento, compone in un certo senso la controversia fra D. Bernoulli e i suoi contemporanei. Possiamo infatti affermare che lo sviluppo in una serie di Fourier convergente non implica di per sé l'analiticità; la funzione sarà però analitica se lo scostamento dalle somme parziali è maggiorato dai termini di una progressione geometrica decrescente.

Confrontando queste stime con quelle della migliore approssimazione mediante polinomi trigonometrici, si vede che per le funzioni regolari le somme di Fourier danno approssimazioni molto buone, vicine alle migliori possibili. Per le funzioni continue non regolari la situazione è peggiore: ne esistono, per esempio, alcune la cui serie di Fourier diverge in tutti i punti razionali.

All'epoca in cui fu scritto questo capitolo rimaneva aperta una questione posta da molto tempo: esiste una funzione continua periodica $f(x)$ la cui serie di Fourier non converga alla funzione in alcun punto quando $n \to \infty$? Il risultato più avanzato ricordato dall'autore era quello di A. N. Kolmogorov, che nel 1926 aveva dimostrato l'esistenza di una funzione periodica integrabile secondo Lebesgue la cui serie di Fourier diverge in ogni punto. Una funzione integrabile secondo Lebesgue può però essere discontinua, come accade nell'esempio di Kolmogorov. La conclusione dell'autore, secondo cui il problema attendeva ancora una soluzione definitiva, va letta nel suo contesto storico.

Per approssimare funzioni continue periodiche arbitrarie mediante polinomi trigonometrici si ricorre ai metodi di sommazione delle serie di Fourier. Al posto delle somme parziali si considerano opportune loro modificazioni. Un metodo particolarmente semplice fu proposto dal matematico ungherese Fejér. Si costruisce formalmente la serie di Fourier della funzione continua periodica, anche se la serie può divergere, e si prendono le medie aritmetiche delle somme parziali:

$$\sigma_n(x) = \frac{S_0(x) + S_1(x) + \cdots + S_n(x)}{n+1}. \qquad (27)$$

Questa è la somma di Fejér di ordine n associata a $f(x)$. Fejér dimostrò che essa converge uniformemente a $f(x)$ per $n \to \infty$.

§8. Approssimazione in media quadratica

Torniamo alla corda vibrante. Supponiamo che, in un certo istante t_0, la corda abbia la forma $y = f(x)$. Si può dimostrare che la sua energia potenziale W, cioè il lavoro disponibile nel ritorno alla posizione di equilibrio, è, per piccole deformazioni e a meno di un fattore costante, uguale all'integrale $W = \int_0^l f'^2(x)\,dx$. Vogliamo ora approssimare $f(x)$ mediante una funzione $\phi(x)$. Accanto alla corda data consideriamone una seconda, di forma $\phi(x)$, e una terza, di forma $f(x) - \phi(x)$. Se l'energia della terza corda,

$$\int_0^l [f'(x) - \phi'(x)]^2\,dx \qquad (28)$$

è piccola, anche la differenza fra le energie delle prime due corde sarà piccola, mantenendo fissata la prima.[10] Se vogliamo dunque che l'energia della seconda differisca poco da quella della prima, dobbiamo

cercare una funzione $\phi'(x)$ che renda quanto più piccolo possibile l'integrale (28). Siamo così condotti al problema di approssimare una funzione - qui $f'(x)$ - in media quadratica.

Nel caso generale il problema si formula così. Sull'intervallo $[a, b]$ sono date una funzione $F(x)$ e una famiglia di funzioni

$$\Phi(x; \alpha_0, \alpha_1, \cdots, \alpha_n), \qquad (29)$$

dipendenti, oltre che da x, dai parametri $\alpha_0, \alpha_1, \ldots, \alpha_n$. Si devono scegliere questi parametri in modo da rendere minimo l'integrale

$$\int_a^b [F(x) - \Phi(x;\ \alpha_0, \alpha_1, \cdots, \alpha_n)]^2\,dx. \qquad (30)$$

L'idea è molto vicina a quella del problema di Čebyšev. Anche qui cerchiamo la migliore approssimazione di $F(x)$ mediante funzioni della famiglia (29), ma secondo il criterio della media quadratica. Non occorre più che $F - \Phi$ sia piccola in ogni punto di $[a, b]$: in una piccola parte dell'intervallo la differenza può essere anche grande, purché l'integrale (30) resti piccolo, come per i due grafici della figura 14. Un piccolo valore della (30) indica che le funzioni F e Φ sono vicine nella maggior parte dell'intervallo.[11] La scelta pratica del criterio d'approssimazione dipende dallo scopo. Nell'esempio della corda è naturale approssimare $f'(x)$ in media quadratica. Questo criterio non era invece adatto ai problemi meccanici di Čebyšev: un componente che oltrepassasse i limiti di tolleranza sarebbe inaccettabile anche se

lo facesse soltanto in una zona piccolissima. Una sola sporgenza potrebbe compromettere l'intera macchina. Čebyšev dovette perciò sviluppare un metodo matematico corrispondente al problema concreto.

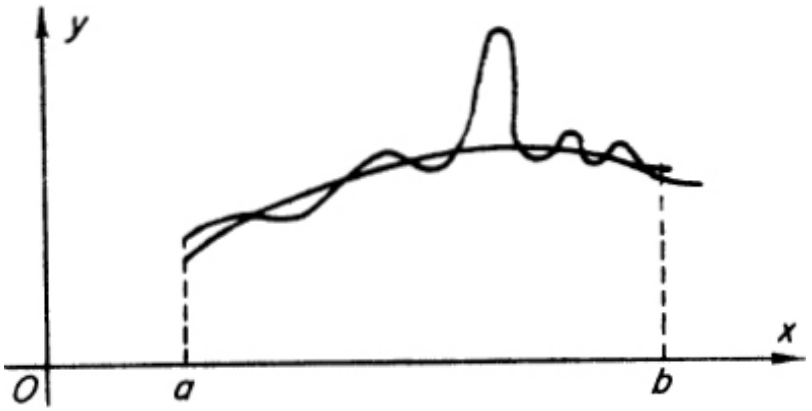


Figura 14.

Dal punto di vista del calcolo, il criterio della media quadratica è più comodo, perché si riconduce a metodi ben sviluppati dell'analisi generale.

Consideriamo, come esempio, il problema caratteristico seguente.

Si vuole ottenere la migliore approssimazione in media quadratica di una funzione continua $f(x)$ su $[a, b]$ mediante somme della forma

$$\sum_1^n \alpha_k \phi_k(x),$$

dove gli α_k sono costanti e le funzioni continue $\phi_k(x)$ costituiscono un sistema ortogonale e normalizzato, cioè ortonormale.

Ciò significa che valgono le uguaglianze

$$\int_a^b \phi_k(x)\,\phi_l(x)\,dx = 0 \text{ per } k \neq l;$$
$$\int_a^b \phi_k^2(x)\,dx = 1 \ (k, l = 1, \ldots, n).$$

Introduciamo i numeri

$$a_k = \int_a^b f(x)\,\phi_k(x)\,dx \quad (k = 1, \ldots, n).$$

Questi numeri a_k si chiamano coefficienti di Fourier di f rispetto al sistema ϕ_k.

Per coefficienti arbitrari α_k, l'ortogonalità e la normalizzazione delle funzioni ϕ_k danno l'identità

$$\int_a^b \left(f - \sum_1^n \alpha_k \phi_k\right)^2 dx = \int_a^b f^2 dx + \sum_1^n \alpha_k^2 - 2\sum_1^n \alpha_k a_k$$

$$= \left(\int_a^b f^2\, dx - \sum_1^n a_k^2\right) + \sum_1^n (\alpha_k - a_k)^2.$$

Il primo termine del secondo membro non dipende dagli α_k. Il minimo si ottiene quindi minimizzando il secondo termine, e ciò avviene precisamente quando ogni α_k coincide con il corrispondente coefficiente di Fourier a_k.

Abbiamo così dimostrato il seguente risultato fondamentale. Se le funzioni ϕ_k formano un sistema ortonormale su $[a, b]$, la somma $\sum_1^n \alpha_k \phi_k(x)$ fornisce la migliore approssimazione in media quadratica di $f(x)$ se e solo se i numeri α_k sono i coefficienti di Fourier di f rispetto alle funzioni $\phi_k(x)$.

Dalle uguaglianze (23) si ricava facilmente che le funzioni

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \cdots$$

costituiscono un sistema ortonormale su $[0, 2\pi]$. Applicato alle funzioni trigonometriche, il risultato precedente assume dunque la forma seguente.

La somma di Fourier $S_n(x)$ di una funzione continua $f(x)$, di periodo 2π, fornisce su $[0, 2\pi]$ la migliore approssimazione in media quadratica fra tutti i polinomi trigonometrici

$$t_n(x) = \alpha_0 + \sum_1^n (\alpha_k \cos kx + \beta_k \sin kx)$$

di ordine non superiore a n.

Da questo risultato e dal teorema di Fejér, enunciato nel §7, discende un altro fatto notevole.

Sia $f(x)$ continua e di periodo 2π, e sia $\sigma_n(x)$ la sua somma di Fejér di ordine n, definita dalla (27).

Introduciamo la notazione

$$\max |f(x) - \sigma_n(x)| = \eta_n .$$

Poiché le somme di Fourier $S_k(x)$, con $k = 0, 1, \ldots, n$, sono polinomi trigonometrici di ordine $k \leqslant n$, anche $\sigma_n(x)$ ha ordine non superiore a n. Per la proprietà di minimo della somma $S_n(x)$ si ha quindi

$$\int_{-\pi}^{\pi} [f(x) - S_n(x)]^2\, dx \leqslant \int_{-\pi}^{\pi} [f(x) - \sigma_n(x)]^2\, dx \leqslant \int_{-\pi}^{\pi} \eta_n^2\, dx = 2\pi\eta_n^2 .$$

Per il teorema di Fejér, η_n tende a zero quando $n \to \infty$. Otteniamo così il risultato seguente.

Per ogni funzione continua di periodo 2π vale l'uguaglianza

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 \, dx = 0.$$

Si dice in questo caso che le somme di Fourier di $f(x)$ convergono alla funzione in media quadratica quando l'ordine n cresce indefinitamente.

L'affermazione vale in realtà per una classe più ampia: quella delle funzioni integrabili secondo Lebesgue insieme al loro quadrato.

Ci fermiamo qui, senza esporre altri risultati interessanti della teoria delle serie di Fourier e dei sistemi ortogonali basati sull'approssimazione in media quadratica. Importanti applicazioni fisiche dei sistemi ortogonali sono già state presentate nel capitolo VI. Questi argomenti saranno ripresi, da un punto di vista in parte diverso, nel capitolo XIX.

Letture consigliate

N. I. Ahiezer, *Theory of approximation*, Frederick Ungar, New York, 1956.

D. Jackson, *The theory of approximation*, American Mathematical Society, Providence, R. I., 1930.

J. L. Walsh, *Interpolation and approximation by rational functions in the complex domain*, 2ª ed., American Mathematical Society, Providence, R. I., 1956.

Note al capitolo XII

[1] Per giustificare completamente l'affermazione occorre verificare che il valore assoluto della differenza $(4x/\pi^2)(\pi - x) - \sin x$ resti inferiore a 0,06 non soltanto per $x = \pi/4$ e $x = \pi/6$, ma per ogni x in $[0, \pi]$. Non svolgiamo qui tale verifica. Il valore 0,05 indicato nell'originale è stato corretto; si veda la nota editoriale.

[2] Uno zero della funzione $f(x)$ è un valore x_k per il quale $f(x_k) = 0$. Per maggiori dettagli sui polinomi di Čebyšev si veda il §5.

[3] Se una quantità α_n, dipendente da $n = 1, 2, \ldots$, soddisfa $|\alpha_n| < C/n^k$, con C indipendente da n, diciamo che è di ordine n^{-k}. Qui l'espressione indica una maggiorazione dell'errore, non necessariamente il suo ordine asintotico esatto.

[4] In questo caso $x_0, x_1, \ldots, x_{n-1}$ sono i punti medi delle parti uguali dell'intervallo $[a, b]$, e non i punti di suddivisione come nelle formule (6) e (9).

[5] Dettagli sui calcoli relativi a questi meccanismi si trovano nel volume II dell'«Eredità scientifica di P. L. Čebyšev», Accademia delle scienze dell'URSS, 1945.

[6] L'autore osserva che i polinomi di Bernštejn, nonostante la semplicità, sono poco usati nel calcolo pratico dell'approssimazione: la loro convergenza può essere lenta anche per funzioni con buone proprietà di derivabilità.

[7] La questione è direttamente collegata all'equazione differenziale della corda vibrante

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \left(a = \frac{l}{\pi}\alpha\right),$$

discussa nel capitolo VI.

[8] La funzione $f(x)$ ha periodo ω se soddisfa $f(x + \omega) = f(x)$.

[9] Il teorema si estende anche a una classe di funzioni discontinue, quelle a variazione limitata. Nei punti di continuità la serie converge al valore della funzione; nei punti di salto converge alla semisomma dei limiti laterali. In presenza di discontinuità, la convergenza della serie non può essere uniforme su tutto il periodo.

[10] Infatti, se

$$\int_0^l f^2(x)\,dx \le M^2 \text{ e } \int_0^l \phi'^2(x)\,dx \le M^2,$$

allora

$$\left| \int_0^l f'^2\,dx - \int_0^l \phi'^2\,dx \right|$$

$$\le \left(\sqrt{\int_0^l f'^2\,dx} + \sqrt{\int_0^l \phi'^2\,dx}\right) \left|\sqrt{\int_0^l f'^2\,dx} - \sqrt{\int_0^l \phi'^2\,dx}\right| \le 2M\sqrt{\int_0^l (f' - \phi')^2\,dx}.$$

[11] Nel capitolo XIX vedremo la profonda analogia fra la vicinanza delle funzioni in media quadratica e la distanza fra punti nello spazio ordinario.

Nota editoriale

La traduzione conserva le otto sezioni, tutte le quattordici figure, le formule, la bibliografia e le undici note del capitolo. Non riproduce la paginazione dell'edizione originale. Le parole inglesi nelle formule sono tradotte e i refusi matematici evidenti sono corretti. I polinomi approssimanti di grado n sono intesi di grado non superiore a n, salvo quando si prescrive un coefficiente principale non nullo.

Nel §2 l'errore massimo dell'approssimazione di sin x mediante $4x(\pi - x)/\pi^2$ è circa 0,05601: il limite 0,05 dell'originale è sostituito con 0,06. Il polinomio di Taylor quadratico centrato in $\pi/2$ vale, in $x = 0$, $1 - \pi^2/8 \approx -0{,}234$. Nella serie geometrica di $1/(1 + x^2)$ è ripristinato il segno meno davanti a x^6; nel resto d'interpolazione e nella sua maggiorazione il denominatore corretto è $(n + 1)!$, anziché $n!$. Si distingue l'interpolazione nei nodi di Čebyšev dai polinomi di Čebyšev stessi, e il teorema di Faber è formulato per uno schema prefissato di nodi.

Nel §3 sono corretti la scrittura f(c) e i valori dell'esempio di Simpson: su ciascun semintervallo si ottiene circa 1,00228 e su $[0, \pi]$ circa 2,00456, con errore circa 0,00456. La formula composta richiede un numero pari di sottointervalli uguali. Gli ordini dell'errore sono maggiorazioni garantite sotto le opportune ipotesi di derivate limitate, non uguaglianze asintotiche per ogni singola funzione. Una maggiore regolarità, da sola, non aumenta in generale l'ordine di una regola a grado di esattezza fissato; restano possibili cancellazioni particolari.

Nei §§4-6 sono precisati l'ordinamento dei punti di alternanza, la continuità della funzione, il caso di errore nullo e la normalizzazione monica dei polinomi che si discostano meno da zero. Nella norma uniforme è ripristinato il valore assoluto. La derivata di $M \cos(n \arccos x)$ ha segno positivo davanti al quoziente; le stime con $\sqrt{(1 - x^2)}$ al denominatore valgono nei punti interni. Sono corretti l'intervallo $[a, b]$, l'espressione $|\alpha - r|$ e il nome Bernštejn. La somma che definisce il polinomio di Bernštejn inizia da $k = 0$. Per Jackson si intende fissato l'intervallo; per la stima geometrica si intende analiticità in un intorno dell'intervallo chiuso. La stima dell'interpolazione è scritta con $\ln(n + 1)$, valida anche per $n = 1$.

Nei §§7-8 sono corretti il periodo 2π, l'indice iniziale $l = 0$ nello sviluppo dell'onda triangolare e la condizione $k \neq l$ per l'ortogonalità. Per m pari, la simmetria usata nel calcolo dei coefficienti è centrale rispetto a $(\pi/2, 0)$. La stima contenente $\ln n$ è intesa per $n \geq 2$. La lettera Φ della famiglia di funzioni è resa coerente con le formule. Si precisa il valore limite della serie nei punti di salto delle funzioni a variazione limitata. Le affermazioni sulla questione allora aperta della divergenza delle serie di Fourier, sulle applicazioni e sullo stato delle ricerche sono riportate come testimonianza storica dell'epoca dell'opera, non come aggiornamento dello stato attuale della matematica.

CAPITOLO XIII

Metodi di approssimazione e tecniche di calcolo

V. I. Krylov

§1. Approssimazione e metodi numerici

Caratteristiche dei metodi di approssimazione.

In molti casi l'applicazione della matematica allo studio dei fenomeni del mondo esterno si fonda sul carattere quantitativo delle leggi che li governano: queste possono essere descritte mediante formule, equazioni o disuguaglianze. Diventa così possibile studiare numericamente i fenomeni ed eseguire i calcoli tanto necessari nella vita pratica.

Una volta individuata una legge quantitativa, la si può studiare con metodi puramente matematici. Per fissare le idee consideriamo una legge espressa da un'equazione: per esempio il moto di un corpo nella meccanica newtoniana, la conduzione del calore o la propagazione delle oscillazioni elettromagnetiche. Equazioni di questo genere sono esaminate nei capitoli V e VI. Di solito all'equazione si aggiungono condizioni che la soluzione deve soddisfare - nei capitoli citati, condizioni iniziali e al contorno - e che, nei problemi ben posti qui considerati, determinano un'unica soluzione.

I primi e più importanti compiti matematici saranno i seguenti.

1. Dimostrare l'esistenza di una soluzione. Anche quando dal punto di vista fisico sembra evidente che il problema sia risolubile, la dimostrazione matematica dell'esistenza di una soluzione per il problema rigorosamente formulato costituisce di solito una verifica necessaria dell'adeguatezza della formulazione. Per un'ampia classe di problemi è possibile fornire tale dimostrazione.

2. Cercare un'espressione esplicita, una formula, per la grandezza che caratterizza il fenomeno. Di solito ciò riesce soltanto nei casi più semplici. Spesso, inoltre, l'espressione ottenuta è così complicata che ricavarne i risultati numerici desiderati è molto difficile o addirittura impossibile in pratica.

3. Trovare un procedimento per costruire una formula approssimata che dia la soluzione con qualsiasi grado di precisione desiderato. Questo è possibile in molti casi.

4. Molto spesso si possono trovare uno o più metodi per calcolare direttamente la soluzione in forma numerica.

Lo sviluppo di questi metodi numerici, molti dei quali approssimati, per risolvere problemi scientifici e tecnici ha dato origine a un particolare ramo della matematica, solitamente chiamato matematica del calcolo o analisi numerica.

I metodi del calcolo numerico sono per loro natura approssimati, poiché ogni grandezza viene calcolata soltanto con un numero finito di cifre: per esempio fino alla quinta o alla sesta cifra decimale.

Per le applicazioni ciò è sufficiente: spesso non occorre conoscere il valore esatto di una grandezza. Nei problemi tecnici, per esempio, il risultato serve abitualmente a stabilire le dimensioni o altri parametri di un manufatto. Ogni lavorazione è soggetta ad approssimazioni; una precisione di calcolo che superi inutilmente quella imposta dalle tolleranze di fabbricazione non offre quindi un vantaggio pratico.

Per il calcolo non è dunque sempre necessario disporre di formule esatte o di soluzioni esatte delle equazioni. Si possono sostituire formule ed equazioni con altre, dichiaratamente approssimate, purché abbastanza vicine alle originali da mantenere l'errore risultante entro limiti assegnati.

Torneremo su questa sostituzione di un problema con un altro. Per ora vogliamo sottolineare la prima caratteristica dei metodi numerici: di regola essi possono produrre soltanto risultati approssimati, che sono però proprio quelli richiesti dalla pratica.

Consideriamo ora un secondo aspetto. In ogni calcolo possiamo operare soltanto su un numero finito di cifre e dobbiamo ottenere il risultato mediante un numero finito di operazioni aritmetiche. Se usiamo una formula, occorre prima trasformarla in modo che contenga soltanto un numero finito di termini e di parametri. È noto, per esempio, che molte funzioni si possono rappresentare come somma di una serie di potenze:

$$f(x) = c_0 + c_1 x + c_2 x^2 + \cdots . \qquad (1)$$

La funzione sin x, dove x misura un angolo in radianti, ammette lo sviluppo

$$\sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots .$$

Per trovare il valore esatto di $f(x)$ occorrerebbe sommare «tutti» i termini della serie (1), cosa in generale impossibile con un calcolo finito. Per ottenere un valore approssimato basta invece prendere un numero finito di termini. Si può dimostrare, per esempio, che per calcolare sin x con errore inferiore a $4 \cdot 10^{-5}$, per angoli da zero a metà di un angolo retto, basta fermarsi al termine in x^5. Si sostituisce così sin x con il polinomio

$$\frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} .$$

Quando un problema di analisi richiede di determinare una funzione, per risolverlo numericamente dobbiamo sostituirlo, in un modo o nell'altro, con la ricerca di un certo numero di parametri numerici. La conoscenza di questi parametri deve permettere di

calcolare approssimativamente la funzione incognita. Illustriamo il procedimento con un esempio.

Si debba risolvere, nell'intervallo $a \leqslant x \leqslant b$, il problema al contorno per l'equazione differenziale

$$L(y) - f(x) = y'' + p(x)\,y' + q(x)\,y - f(x) = 0 \qquad (2)$$

con condizioni $y(a) = 0$ e $y(b) = 0$. Uno dei procedimenti possibili è il metodo di Galerkin. Si parte da un sistema di funzioni linearmente indipendenti $\omega_1(x)$, $\omega_2(x)$, ..., che soddisfano le condizioni al contorno; si veda il capitolo VI, §5. Il sistema viene scelto completo: nell'opportuno spazio di funzioni integrabili, una funzione ortogonale a tutte le ω_k, con $k = 1, 2, \ldots$, deve essere nulla quasi ovunque nell'intervallo. La condizione che $y(x)$ soddisfi l'equazione differenziale (2) si può allora esprimere mediante le condizioni di ortogonalità

$$\int_a^b [L(y) - f]\omega_k\,dx = 0 \qquad (k = 1, 2, \cdots). \qquad (3)$$

Supponiamo che la soluzione del problema si possa sviluppare in una serie delle funzioni ω_k:

$$y(x) = a_1\omega_1(x) + a_2\omega_2(x) + \cdots. \qquad (4)$$

Cerchiamo le condizioni che i coefficienti a_k devono soddisfare. Sotto le opportune ipotesi di convergenza, la somma della serie (4) soddisfa le condizioni al contorno, perché le soddisfa ciascuna funzione del sistema. Resta da scegliere i coefficienti in modo che valgano le (3). I coefficienti sono però infiniti, e in generale non è possibile calcolarli tutti. Per semplificare, conserviamo soltanto un numero finito di termini della (4), ottenendo

$$y(x) \approx a_1\omega_1(x) + \cdots + a_n\omega_n(x). \qquad (5)$$

Disponendo di soli n parametri liberi a_k, non possiamo in generale soddisfare le (3) per tutte le funzioni ω_k. Dobbiamo quindi rinunciare alla soluzione esatta dell'equazione differenziale (2). È però naturale attendersi che la somma (5) soddisfi l'equazione con un piccolo errore, se n è abbastanza grande e si impongono le (3) per le prime n funzioni del sistema. La validità di questa aspettativa deve naturalmente essere giustificata nelle ipotesi del problema. Si arriva così alle equazioni del metodo di Galerkin:

$$\int_a^b \left[L\left(\sum_{k=1}^{n} a_k\omega_k\right) - f\right]\omega_i\,dx = 0 \qquad (i = 1, 2, \cdots, n).$$

Determinati i coefficienti a_k, si costruisce mediante la (5) un'espressione approssimata della funzione cercata.

Una semplificazione analoga si incontra nella soluzione dei problemi variazionali con il metodo di Ritz, nell'analisi armonica approssimata delle funzioni e in molti altri problemi.

Consideriamo un altro esempio di semplificazione. Si voglia determinare una funzione y di una o più variabili risolvendo un'equazione funzionale, per esempio differenziale o integrale.

Come parametri che descrivono la funzione scegliamo i suoi valori y_1, y_2, ..., y_n in un certo sistema di punti, cioè su una griglia.

L'equazione funzionale deve allora essere sostituita da un sistema di equazioni numeriche con n incognite y_k. Di regola la sostituzione si può fare in molti modi. In ogni caso occorre assicurarsi che la soluzione del sistema numerico differisca abbastanza poco da quella dell'equazione funzionale.

Ecco alcuni esempi. Quando risolviamo l'equazione differenziale del primo ordine $y' = f(x, y)$ con il metodo di Eulero, la sostituiamo con uno schema ricorrente che permette di calcolare approssimativamente ogni valore successivo della funzione a partire dal precedente; si veda il capitolo V, §5:

$$y_{n+1} = y_n + (x_{n+1} - x_n) f(x_n , y_n).$$

Per risolvere approssimativamente l'equazione di Laplace

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

con il metodo delle differenze finite su una griglia, si sostituisce l'equazione con un sistema algebrico lineare; si veda il capitolo VI, §5:

$$u(x + h, y) + u(x, y + h) + u(x - h, y) + u(x, y - h) - 4u(x, y) = 0.$$

Consideriamo ancora un esempio. Si debba risolvere numericamente l'equazione integrale

$$y(x) = f(x) + \int_a^b K(x, s)\, y(s)\, ds. \qquad (6)$$

Indichiamo con x_1, x_2, ..., x_n i punti in cui vogliamo trovare i valori della funzione incognita. Per costruire il sistema numerico che sostituisce la (6), richiediamo che essa sia soddisfatta non per tutti gli x dell'intervallo $a \leqslant x \leqslant b$, ma soltanto nei nodi x_i, con $i = 1, 2, ..., n$:

$$y(x_i) = f(x_i) + \int_a^b K(x_i , s)\, y(s)\, ds.$$

Sostituiamo poi l'integrale con una formula di quadratura approssimata - dei trapezi, di Simpson o un'altra[1] - avente nodi x_1, ..., x_n:

$$\int_a^b K(x_i , s)\, y(s)\, ds \approx \sum_{j=1}^{n} A_{ij} K(x_i , x_j)\, y(x_j).$$

Per determinare i valori cercati $y(x_i)$ otteniamo il sistema di equazioni algebriche lineari

$$y(x_i) = f(x_i) + \sum_{j=1}^{n} A_{ij} K(x_i, x_j) y(x_j),$$
$$(i = 1, 2, ..., n). \quad (7)$$

In tutti i metodi esaminati, la ricerca di una funzione incognita si è ricondotta alla determinazione di parametri che la descrivono approssimativamente. La precisione dipende quindi da quanto bene il sistema di parametri rappresenta la funzione: per esempio da quanto bene essa si lascia approssimare mediante un'espressione della forma (5), oppure ricostruire dai suoi valori in determinati punti. Queste questioni costituiscono un ramo specifico della matematica, la teoria dell'approssimazione delle funzioni, esaminata nel capitolo XII. Si comprende così la sua grande importanza per la matematica applicata.

Convergenza dei metodi approssimati e stima dell'errore.

Esaminiamo più da vicino i requisiti di un metodo numerico. Il più semplice e fondamentale è la possibilità di trovare la grandezza cercata con qualsiasi grado di precisione assegnato.

La precisione richiesta può variare molto da un problema all'altro. Nelle valutazioni tecniche grossolane dell'epoca potevano bastare due o tre cifre decimali, mentre molti calcoli ingegneristici venivano eseguiti con tre o quattro. Nei calcoli scientifici si richiede spesso una precisione molto maggiore. In generale, l'esigenza di una maggiore accuratezza è cresciuta nel corso del tempo.

Sono perciò particolarmente importanti i metodi e i procedimenti che permettono di ottenere risultati con precisione arbitrariamente elevata. Questi metodi si chiamano convergenti. Poiché ricorrono frequentemente nella pratica e i loro requisiti sono rappresentativi, ci riferiremo a essi nel seguito.

Sia x il valore esatto della grandezza cercata. A ogni metodo di questo tipo corrisponde una successione di approssimazioni $x_1, x_2, \ldots, x_n, \ldots$ della soluzione x.

Una volta stabilita la costruzione delle approssimazioni, il primo problema teorico è dimostrarne la convergenza, $x_n \to x$; se il metodo non converge sempre, occorre precisare le condizioni nelle quali converge.

Dimostrata la convergenza, si pone il problema più difficile e delicato di stimarne la rapidità, cioè di valutare quanto velocemente x_n tende a x per $n \to \infty$. In teoria un metodo convergente garantisce qualsiasi precisione desiderata, purché si scelga un indice n abbastanza grande. Di regola, però, al crescere di n cresce anche il lavoro necessario a calcolare x_n. Se la convergenza è lenta, raggiungere la precisione richiesta può comportare una quantità enorme di calcoli.

Nella matematica e soprattutto nelle sue applicazioni sono noti molti procedimenti convergenti che richiedono una mole di calcolo superiore alle possibilità pratiche, anche usando le macchine veloci disponibili all'epoca dell'opera.[2]

Una convergenza troppo lenta è uno dei criteri con cui si giudicano gli svantaggi di un metodo. Non è però l'unico: nel confronto occorre considerare molti altri aspetti, in particolare la facilità di esecuzione sulle macchine. Fra due metodi può essere preferibile quello un po' più lento a convergere, se i suoi calcoli risultano più agevoli da realizzare su una macchina.

L'errore introdotto sostituendo x con x_n è la differenza $x - x_n$. Il suo valore esatto è sconosciuto. Per valutare la rapidità di convergenza cerchiamo quindi un limite superiore del suo valore assoluto, cioè una quantità A_n tale che

$$| x - x_n | \leqslant A_n ,$$

che chiameremo stima dell'errore. Più avanti ne daremo alcuni esempi. Di solito si giudica la rapidità di convergenza osservando quanto rapidamente A_n diminuisce al crescere di n. Perché la stima rifletta bene l'effettiva vicinanza fra x_n e x, essa dovrebbe differire poco da $|x - x_n|$. Deve inoltre essere effettiva, ossia calcolabile: altrimenti non la si può utilizzare concretamente.

Sia x una grandezza numerica da determinare risolvendo un'equazione. Supponiamo di poter scrivere l'equazione nella forma

$$x = \phi(x). \qquad (8)$$

Applichiamo il metodo di iterazione, detto anche delle approssimazioni successive. Per illustrarlo e ricavarne le stime consideriamo una sola equazione numerica, benché il metodo si applichi anche a sistemi, equazioni differenziali, equazioni integrali e molti altri problemi. L'applicazione alle equazioni differenziali ordinarie è già stata illustrata nel capitolo V, §5.

Supponiamo di avere trovato in qualche modo un valore approssimato x_0 di una radice. Se fosse una soluzione esatta della (8), sostituendolo nel secondo membro $\phi(x)$ otterremmo ancora x_0. Poiché in generale non lo è, il risultato sarà diverso. Indichiamolo con $x_1 = \phi(x_0)$.

Per stabilire quando x_1 sia più vicino di x_0 alla soluzione esatta, ricorriamo a un'interpretazione geometrica. Consideriamo la funzione

$$y = \phi(x). \qquad (9)$$

Rappresentiamo i numeri x e y mediante punti di un asse numerico. La (9) associa a ogni punto x un punto y dello stesso asse. Essa definisce dunque una trasformazione dell'asse numerico in sé.

Consideriamo il segmento di estremi x_1 e x_2. La trasformazione (9) manda i suoi estremi nei punti

$$y_1 = \phi(x_1) \text{ e } y_2 = \phi(x_2).$$

Agli estremi del segmento iniziale corrispondono quindi gli estremi del segmento che congiunge y_1 e y_2. Il rapporto

$$k = \frac{| y_2 - y_1 |}{| x_2 - x_1 |}$$

si chiama coefficiente di dilatazione relativo a questa coppia di punti. Se $k < 1$, la distanza fra i due punti si contrae.

Torniamo alla (8). Essa richiede che il punto cercato venga trasformato in sé stesso dalla (9). Risolvere l'equazione equivale dunque a trovare un punto dell'asse numerico che rimanga fisso sotto la trasformazione.

Consideriamo ora il segmento di estremi x e x_0, il primo dei quali è il punto fisso. La trasformazione lascia fermo x e manda x_0 in x_1. Se ϕ contrae la distanza fra ogni coppia di punti, x_1 sarà certamente più vicino di x_0 alla radice della (8).

Per ottenere approssimazioni convergenti alla soluzione esatta, ripetiamo la trasformazione e costruiamo la successione

$$x_1 = \phi(x_0), \quad x_2 = \phi(x_1), \cdots, x_{n+1} = \phi(x_n), \cdots . \qquad (10)$$

Dimostriamo ora la convergenza di questa successione, sotto le ipotesi che seguono.[3]

Supponiamo che $\phi(x)$ sia definita su un intervallo $[a, b]$ e che la (9) trasformi questo intervallo in sé: per ogni x appartenente a $[a, b]$, anche $\phi(x)$ vi appartiene. Supponiamo inoltre che x_0 sia nell'intervallo; allora vi rimarranno tutte le approssimazioni successive. Vale il seguente teorema. Se ϕ è continua sull'intervallo e derivabile al suo interno, con derivata che soddisfa

$$|\phi'| \leqslant q < 1$$

per una costante $0 \leq q < 1$, l'equazione (8) possiede una radice x^* in $[a, b]$. La successione (10) converge a questa radice e la rapidità di convergenza è caratterizzata dalla stima

$$|x^* - x_n| \leqslant \frac{m}{1-q} q^n,$$

dove $m = |x_0 - \phi(x_0)| = |x_0 - x_1|$. La radice della (8) nell'intervallo è unica.

Per dimostrare queste affermazioni stimiamo la differenza $x_2 - x_1$. La formula di Taylor al primo ordine, nella forma del teorema del valor medio, dà

$$x_2 - x_1 = \phi(x_1) - \phi(x_0) = \phi'(\xi_0)(x_1 - x_0).$$

dove ξ_0 è compreso fra x_1 e x_0, e appartiene quindi a $[a, b]$. Pertanto $|\phi'(\xi_0)| \leqslant q$ e

$$|x_2 - x_1| \leqslant q\,|x_1 - x_0| = mq.$$

Analogamente

$$|x_3 - x_2| = |\phi(x_2) - \phi(x_1)| = |\phi'(\xi_1)(x_2 - x_1)| \leq q|x_2 - x_1| \leq mq^2.$$

Proseguendo, otteniamo per ogni n la disuguaglianza

$$|x_{n+1} - x_n| \leqslant mq^n. \qquad (11)$$

Dimostriamo ora che la successione x_n converge. Consideriamo la serie ausiliaria

$$x_0 + (x_1 - x_0) + (x_2 - x_1) + \cdots + (x_n - x_{n-1}) + \cdots . \qquad (12)$$

La somma dei primi $n + 1$ termini è

$$s_{n+1} = x_0 + (x_1 - x_0) + \cdots + (x_n - x_{n-1}) = x_n .$$

Quindi il limite di S_{n+1} coincide con quello di x_n; l'esistenza di un limite finito della successione equivale alla convergenza della serie (12). Confrontiamo quest'ultima con la serie

$$| x_0 | + m + mq + \cdots + mq^{n-1} + \cdots .$$

Per la stima (11), i valori assoluti dei termini della (12) non superano i corrispondenti termini di questa serie. A parte il primo termine $|x_0|$, essa è una serie geometrica di ragione q; poiché $q < 1$, converge. Converge dunque anche la (12), e la successione (10) ammette un limite finito x^*:

$$\lim_{n\to\infty} x_n = x^* .$$

Il limite appartiene a $[a, b]$, perché tutti gli x_n appartengono a questo intervallo chiuso.

Passando al limite nell'uguaglianza $x_{n+1} = \phi(x_n)$ e usando la continuità di ϕ, otteniamo $x^* = \phi(x^*)$. Il limite soddisfa quindi la (8). Stimiamo ora la distanza di x_n da x^*. Confrontiamo x_n con una qualsiasi approssimazione successiva x_{n+p}:

$$\begin{aligned} | x_{n+p} - x_n | &= | (x_{n+p} - x_{n+p-1}) + (x_{n+p-1} - x_{n+p-2}) + \cdots + (x_{n+1} - x_n) | \\ &\leqslant mq^{n+p-1} + mq^{n+p-2} + \cdots + mq^n \\ &= \frac{mq^n - mq^{p+n}}{1 - q} . \end{aligned}$$

Facendo tendere p all'infinito, poiché $x_{n+p} \to x^*$ e $q^{n+p} \to 0$, ricaviamo

$$| x^* - x_n | \leqslant \frac{m}{1 - q} q^n .$$

Resta l'unicità. Sia x' un'altra soluzione nell'intervallo. Allora

$$| x' - x^* | = | \phi(x') - \phi(x^*) | = | \phi'(\xi)(x' - x^*) | \leqslant q | x' - x^* |,$$

da cui

$$(1 - q) | x' - x^* | \leqslant 0.$$

Poiché $1 - q > 0$, la disuguaglianza può valere soltanto se $|x' - x^*| = 0$, cioè se le due soluzioni coincidono.

Il teorema fornisce condizioni sufficienti per la convergenza e permette di stimare il numero di passi necessari: indica quanto grande debba essere n perché la sostituzione di x^* con x_n dia

la precisione richiesta. La stima è effettiva, perché le quantità m e q nella disuguaglianza $|x^* - x_n| \le [m/(1-q)]q^n$ si possono determinare studiando la funzione ϕ.

Consideriamo, per esempio, l'equazione $x = k \tan x$, che ricorre in molte applicazioni. Prendiamo $k = 0{,}5$ e cerchiamo la più piccola radice positiva di $x = \frac{1}{2}\tan x$. Una tavola o il grafico della tangente mostrano facilmente che essa è vicina a 1 e un po' maggiore di 1.

Per ottenere la condizione $|\phi'| \leqslant q < 1$, richiesta dal teorema, invertiamo la tangente sul ramo che contiene la radice cercata. Otteniamo l'equazione equivalente, per questa radice, $x = \arctan(2x)$. Per esempio, sull'intervallo [1, 1,2] questa funzione manda l'intervallo in sé e la sua derivata non supera 0,4.

Riportiamo i risultati del calcolo. Come approssimazione iniziale scegliamo $x_0 = 1$. Le successive approssimazioni si ricavano da una tavola dell'arcotangente:

$$
\begin{aligned}
x_1 &= \arctan 2 = 1.10715,\\
x_2 &= \arctan 2.21430 = 1.14660,\\
x_3 &= \arctan 2.29320 = 1.15959,\\
x_4 &= \arctan 2.31918 = 1.16370,\\
x_5 &= \arctan 2.32740 = 1.16498,\\
x_6 &= \arctan 2.32996 = 1.16538,\\
x_7 &= \arctan 2.33076 = 1.16550,\\
x_8 &= \arctan 2.33100 = 1.16554,\\
x_9 &= \arctan 2.33108 = 1.16555,\\
x_{10} &= \arctan 2.33110 = 1.16556,\\
x_{11} &= \arctan 2.33112 = 1.16556.
\end{aligned}
$$

Alla precisione riportata, il calcolo può fermarsi: le iterazioni successive ripetono il medesimo valore arrotondato,

$$x^* \approx 1{,}16556.$$

La figura 1 illustra geometricamente le approssimazioni. La convergenza è così rapida che nel disegno x_4 è già indistinguibile da x^*.

Diamo un altro esempio del metodo di iterazione. Risolviamo numericamente l'equazione integrale

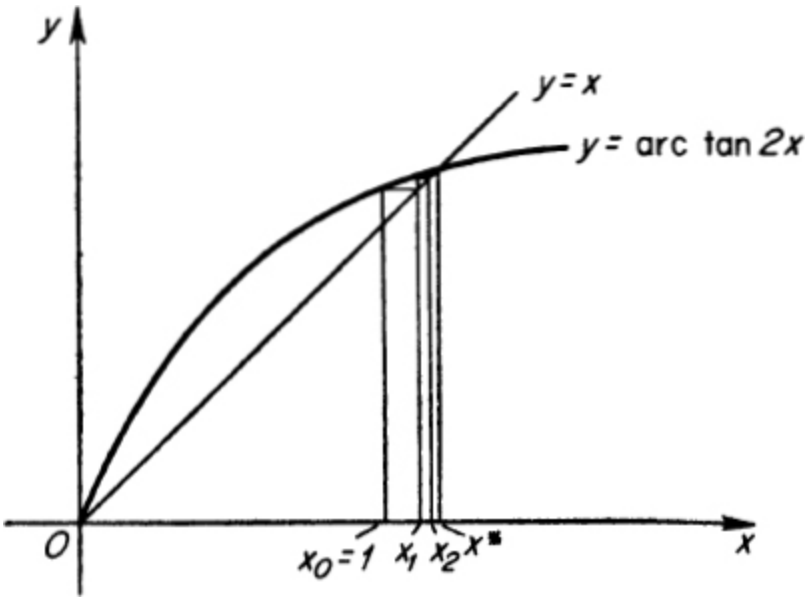


Figura 1.

$$y(x) = \frac{1}{6}\int_0^1 e^{xt}y(t)\,dt + e^x - \frac{1}{6}\frac{1}{x+1}(e^{x+1}-1). \quad (13)$$

La sua soluzione esatta è $y = e^x$.

Sostituiamo dapprima l'equazione integrale con un sistema algebrico lineare. Dividiamo l'intervallo d'integrazione [0, 1] in quattro parti uguali, mediante i punti $t = 0,\ \frac{1}{4},\ \frac{1}{2},\ \frac{3}{4},\ 1$. Indichiamo con y_0, y_1, y_2, y_3, y_4 i valori dell'incognita in questi punti. Richiediamo che l'equazione sia soddisfatta per $x_0 = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$, sostituendo l'integrale con la somma di Simpson su quattro sottointervalli; si veda il capitolo XII, §3, formula (6). Otteniamo il sistema

$$\begin{aligned}
y_0 &= \tfrac{1}{6}(0.083333y_0 + 0.333333y_1 + 0.166667y_2 \\
&\qquad + 0.333333y_3 + 0.083333y_4) + 0.713619, \\
y_1 &= \tfrac{1}{6}(0.083333y_0 + 0.354831y_1 + 0.188858y_2 \\
&\qquad + 0.402077y_3 + 0.107002y_4) + 0.951980, \\
y_2 &= \tfrac{1}{6}(0.083333y_0 + 0.377716y_1 + 0.214004y_2 \\
&\qquad + 0.484997y_3 + 0.137393y_4) + 1.261867, \\
y_3 &= \tfrac{1}{6}(0.083333y_0 + 0.402077y_1 + 0.242499y_2 \\
&\qquad + 0.585018y_3 + 0.176417y_4) + 1.664181, \\
y_4 &= \tfrac{1}{6}(0.083333y_0 + 0.428008y_1 + 0.274787y_2 \\
&\qquad + 0.705667y_3 + 0.226523y_4) + 2.185861.
\end{aligned}$$

Risolviamo questo sistema per iterazione. Come approssimazioni iniziali assumiamo i termini noti delle rispettive equazioni: $y_0^{(0)} = 0.713619$, $y_1^{(0)} = 0.951980$, e così via. I valori delle approssimazioni successive sono riportati nella tavola 1.

Tavola 1.

Iterazione	y_0	y_1	y_2	y_3	y_4
1	0,93428	1,20841	1,56129	2,01542	2,59972
2	0,98517	1,26699	1,62905	2,09419	2,69173
3	0,99667	1,28021	1,64433	2,11194	2,71245
4	0,99926	1,28319	1,64778	2,11595	2,71713
5	0,99985	1,28386	1,64856	2,11685	2,71818
6	0,99998	1,28402	1,64873	2,11705	2,71842
7	1,00001	1,28405	1,64877	2,11710	2,71847
Soluzione esatta	1,00000	1,28403	1,64872	2,11700	2,71828

L'ultima riga riporta, per confronto, i valori della soluzione esatta. Ulteriori iterazioni non eliminerebbero la differenza fra la soluzione del sistema discreto e quella dell'equazione integrale. Le divergenze nelle ultime cifre dipendono dall'approssimazione introdotta con la somma di Simpson e dalla precisione dei calcoli riportati.

Stabilità dei metodi approssimati.

Le esigenze del calcolo pratico impongono un altro requisito fondamentale: la stabilità del procedimento. Ogni metodo conduce a uno schema numerico e, per ottenere tutti i valori richiesti, occorre spesso eseguire una lunga successione di passi. A ogni passo si calcola soltanto con un certo numero di cifre, introducendo quindi un piccolo errore. Tutti questi errori influenzano il risultato finale.

Uno schema può essere tanto sfavorevole che piccoli errori iniziali esercitino un'influenza sempre maggiore man mano che il calcolo prosegue, causando alla fine grandi scostamenti dai valori esatti.

Consideriamo la soluzione numerica dell'equazione differenziale

$$y' = f(x, y)$$

con condizione iniziale $y(x_0) = y_0$. Si devono trovare i valori della soluzione nei punti equidistanti $x_k = x_0 + kh$, con $k = 0, 1, \ldots$.

Supponiamo di avere già eseguito il calcolo fino al passo n, ottenendo i risultati schematizzati nella tavola 2.

Tavola 2.

x	y	$y' = f$
x_0	y_0	y'_0
x_1	y_1	y'_1
...	...	...
x_{n-1}	y_{n-1}	y'_{n-1}
x_n	y_n	y'_n

Dobbiamo ora determinare y_{n+1}. Il metodo delle poligonali di Eulero dà

$$y_{n+1} = y_n + hy'_n . \qquad (14)$$

Qui y_{n+1} viene calcolato soltanto dai numeri y_n e y'_n, che compaiono nell'ultima riga della tavola. Se volessimo aumentare la precisione usando tutti i dati delle ultime due righe, potremmo costruire la formula

$$y_{n+1} = -4y_n + 5y_{n-1} + h(4y'_n + 2y'_{n-1}). \qquad (15)$$

Se il calcolo fosse eseguito senza errori di arrotondamento, la (14) sarebbe esatta per una soluzione y polinomiale di primo grado, mentre la (15) sarebbe esatta per ogni polinomio di grado non superiore a tre, partendo da dati precedenti esatti. A prima vista sembrerebbe quindi che la (15) debba dare risultati più precisi. Ma essa è inadatta al calcolo ripetuto, perché può amplificare rapidamente gli errori.

I valori delle derivate y'_n e y'_{n-1} sono moltiplicati dal piccolo fattore h; i loro errori entrano quindi con un peso minore rispetto a quelli di y_n e y_{n-1}. Per semplicità supponiamo esatti i valori delle derivate e consideriamo soltanto gli altri errori. Se nel calcolo di y_{n-1} commettiamo un errore $+\varepsilon$ e in quello di y_n un errore $-\varepsilon$, la (15) produce in y_{n+1} un errore $+9\varepsilon$, e in y_{n+2} un errore -41ε. Proseguendo, l'errore cresce rapidamente. La formula conduce dunque a un procedimento instabile, che deve essere scartato.

L'esempio seguente mostra quanto uno schema instabile possa alterare i risultati. Risolviamo $y' = y$ con condizione iniziale $y(0) = 1$; la soluzione esatta è $y = e^x$. Scegliamo il passo $h = 0{,}01$, quindi $x_k = 0{,}01k$. Calcoliamo la soluzione approssimata sia con il metodo di Eulero (14), sia con la formula (15). Per confronto, la tavola 3 contiene i valori della soluzione esatta con sette cifre decimali.

Nei primi passi i risultati della (15) sono più precisi di quelli del metodo di Eulero. Dopo pochi passi, però, l'instabilità altera fortemente i valori approssimati e porta a numeri molto lontani dalla soluzione vera.

Tavola 3.

x	Soluzione esatta	Approssimazione con la (14)	Approssimazione con la (15)
0,00	1,0000000	1,0000000	1,0000000
0,01	1,0100502	1,0100000	1,0100502
0,02	1,0202013	1,0201000	1,0202012
0,03	1,0304545	1,0303010	1,0304553
0,04	1,0408108	1,0406040	1,0408070
0,05	1,0512711	1,0510100	1,0512899
0,06	1,0618365	1,0615201	1,0617431
0,07	1,0725082	1,0721353	1,0729726
0,08	1,0832871	1,0828567	1,0809789
0,09	1,0941743	1,0936853	1,1056460
0,10	1,1051709	1,1046222	1,0481559
0,11	1,1162781	1,1156684	1,3996456
0,12	1,1274969	1,1268250	−0,2808540

Scelta dei metodi di calcolo.

Ogni calcolo numerico qui considerato si può ricondurre, in ultima analisi, alle quattro operazioni di addizione, sottrazione, moltiplicazione e divisione. Descrivere un metodo significa specificare i dati iniziali e prescrivere quali operazioni eseguire e in quale ordine per ottenere il risultato. Un esempio molto semplice mostrerà quanto l'organizzazione dei calcoli dipenda dall'esperienza e dalle conoscenze del matematico che prepara lo schema, e quanto si possa guadagnare scegliendo un metodo adatto al problema.

Si debba risolvere il sistema di n equazioni in n incognite $x_1, x_2, \ldots, x_n$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 , \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 , \\ \cdots\cdots\cdots\

dove Δ è il determinante del sistema,

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

e Δ_j è il determinante ottenuto sostituendo la j-esima colonna di Δ con la colonna dei termini noti.

Supponiamo di usare la formula (16), calcolando i determinanti direttamente dalla loro definizione, senza semplificazioni. Quante moltiplicazioni e divisioni occorreranno? Non conteremo addizioni e sottrazioni, considerate relativamente semplici. Dobbiamo calcolare $n + 1$ determinanti di ordine n. Ognuno contiene $n!$ termini, ciascuno prodotto di n fattori, e richiede quindi $n - 1$ moltiplicazioni per termine. Per tutti i determinanti occorrono $(n + 1)n!(n - 1)$ moltiplicazioni. Aggiungendo le divisioni finali, il totale è $(n^2 - 1)n! + n$.

Scegliamo ora un altro procedimento: l'eliminazione successiva delle incognite, il cui schema è associato al nome di Gauss. Ricaviamo x_1 dalla prima equazione:

$$x_1 = \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2 - \cdots - \frac{a_{1n}}{a_{11}} x_n .$$

Questa operazione richiede n divisioni. La sostituzione nelle altre $n - 1$ equazioni richiede n moltiplicazioni per ciascuna. L'eliminazione di x_1 e la costruzione del sistema ridotto comportano dunque n^2 moltiplicazioni e divisioni. Continuando nello stesso modo, e conservando i coefficienti già normalizzati per la sostituzione all'indietro, si ottiene un totale di $n(2n^2 + 6n - 2)/6$ operazioni di questi due tipi. Si suppongono non nulli i divisori successivi; se necessario si scambiano opportunamente le equazioni. Confrontiamo i risultati: per cinque equazioni, il calcolo diretto dei determinanti richiede 2.885 moltiplicazioni e divisioni, mentre questo schema di eliminazione ne richiede 65.

Per dieci equazioni il primo conteggio dà $(10^2 - 1)10! + 10 \approx 360.000.000$, mentre il secondo dà 430. La mole del lavoro dipende dunque in modo decisivo dal metodo scelto. Una buona organizzazione dello schema di calcolo può ridurre enormemente il numero di operazioni necessarie.

§2. I più semplici strumenti ausiliari di calcolo

Tavole.

Il più antico strumento ausiliario di calcolo è costituito dalle tavole numeriche.[4] Il lettore conosce certamente quelle più semplici: la tavola pitagorica, le tavole dei logaritmi e delle funzioni trigonometriche. Il campo dei problemi affrontabili nella pratica si amplia continuamente; nuovi problemi richiedono nuove formule o introducono nuove funzioni, aumentando così il numero di tavole utili.

Ogni tavola, comunque costruita, contiene risultati di calcoli precedenti e costituisce quindi una sorta di memoria matematica. Le tavole stampate o manoscritte sono destinate alla lettura umana. Si possono però immaginare tavole registrate in forme particolari, per esempio mediante fori su schede, destinate a essere lette dalle macchine. Nell'esposizione dell'autore queste ultime sono considerate ancora poco diffuse e non vengono esaminate qui.

Le tavole più comuni sono quelle dei valori di funzioni. Se y dipende da una sola variabile x, la tavola più semplice ha la forma

x	y
x_1	y_1
x_2	y_2
...	...
x_n	y_n

(17)

e si chiama tavola a una entrata.[5] Da essa si leggono direttamente soltanto i valori corrispondenti agli argomenti tabulati. Per gli altri valori di x occorre ricorrere all'interpolazione, descritta nel capitolo XII.[6] Per facilitarla, le tavole contengono spesso grandezze ausiliarie, di solito differenze prime o seconde. Tavole più specializzate richiedono formule d'interpolazione apposite e riportano i dati necessari a usarle.

Per una funzione di due variabili $u = f(x, y)$, i valori sono disposti in una tavola a doppia entrata della forma

x \ y	y_1	y_2	...	y_m
x_1	u_{11}	u_{12}	...	u_{1m}
x_2	u_{21}	u_{22}	...	u_{2m}
...	...	...	...	...
x_n	u_{n1}	u_{n2}	...	u_{nm}

(17′)

Ogni colonna è a sua volta una tavola a una entrata: la (17′) riunisce quindi molte tavole della forma (17). A parità di numero di valori per ciascuna variabile, una tavola di due variabili è di solito molto più grande di una tavola di una sola variabile. Per questo le funzioni di due variabili vengono tabulate assai meno frequentemente.

Un esempio mostra la rapidità con cui le dimensioni crescono al crescere del numero di variabili. Si voglia tabulare $f(x, y, z, t)$ per 100 valori di ciascuna delle quattro variabili. Supponiamo che bastino tre cifre significative. Per una funzione di una sola variabile, alle stesse condizioni, la tavola conterrebbe cento numeri di tre cifre e starebbe comodamente in una pagina.

Per quattro variabili, invece, vi sono 100^4 combinazioni dei valori di x, y, z, t e altrettanti valori della funzione. L'autore stima che una tavola stampata di queste dimensioni riempirebbe più di 300 volumi.

A causa di tale ingombro, le funzioni di molte variabili vengono raramente tabulate integralmente, salvo casi particolarmente semplici. Negli anni precedenti alla stesura dell'opera era iniziato uno studio sistematico delle classi di funzioni che consentono tavole con un numero di entrate inferiore a quello delle variabili, e dei modi più semplici per costruirle.

Diamo un esempio elementare.

Si voglia tabulare una funzione u di tre variabili x, y, z avente la struttura

$$u = f[\phi(x, y), z].$$

Introducendo la variabile ausiliaria $t = \phi(x, y)$, basta evidentemente costruire due tavole a doppia entrata, considerando la funzione composta

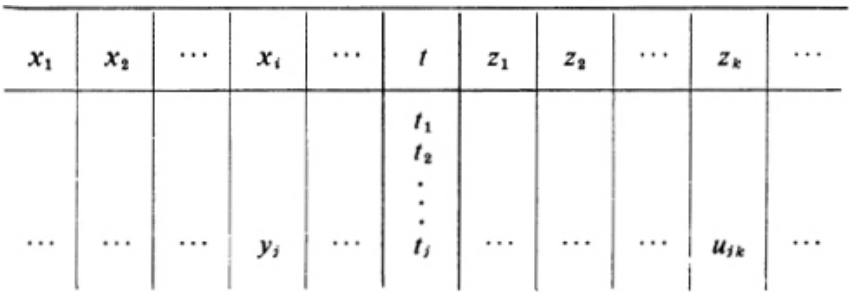
$$u = f(t, z),$$
$$t = \phi(x, y).$$

Per rendere più comodo l'uso delle tavole, possiamo combinarle nel modo seguente. Consideriamo $t = \phi(x, y)$ e, supponendo possibile un'inversione univoca nel campo considerato, risolviamo rispetto a y:

$$y = \Phi(x, t).$$

In linea di principio possiamo tabulare tanto $t = \phi(x, y)$ quanto la funzione inversa $y = \Phi(x, t)$. Qui è più comoda la seconda. Costruiamo due tavole a doppia entrata, per $y = \Phi(x, t)$ e per $u = f(t, z)$, e le disponiamo come nella tavola 4.

Tavola 4.



x_1	x_2	⋯	x_i	⋯	t	z_1	z_2	⋯	z_k	⋯
					t_1					
					t_2					
					⋮					
⋯	⋯	⋯	y_j	⋯	t_j	⋯	⋯	⋯	u_{jk}	⋯

Il valore di u corrispondente ai dati x_i, y_j, z_k si trova così. Nella colonna intestata a x_i cerchiamo il valore y_j, o uno vicino. La stessa riga orizzontale contiene il corrispondente valore di t. Proseguendo lungo la riga fino alla colonna z_k troviamo il valore richiesto $u = f(\phi(x_i, y_j), z_k)$.

In questo esempio, una tavola a tre entrate è quindi sostituita da due tavole a doppia entrata e da una semplice regola di consultazione.

Vari procedimenti di riduzione permettono, in certi casi, di diminuire le dimensioni delle tavole di dieci, cento o perfino mille volte rispetto alla tabulazione completa di tutte le combinazioni delle variabili indipendenti.

Calcolatrici da tavolo.

Quasi altrettanto antichi delle tavole sono i dispositivi di calcolo. Alcuni erano già utilizzati nella Grecia antica.

Nel Seicento Pascal, Morland e Leibniz costruirono i primi modelli di macchine calcolatrici. Da allora questi strumenti furono ripetutamente modificati e perfezionati; il loro uso si diffuse alla fine dell'Ottocento e soprattutto all'inizio del Novecento.

Esamineremo soltanto alcuni tipi di macchine e le possibilità di accelerarne i calcoli. Cominciamo dalle piccole calcolatrici universali da tavolo. Indipendentemente dai dettagli costruttivi, sono progettate per eseguire le quattro operazioni aritmetiche; nei modelli qui descritti, moltiplicazione e divisione si ottengono mediante successioni di addizioni e sottrazioni.

Un modello tipico delle prime macchine di questo genere è l'aritmometro a ruote di Odhner. Per impostare un numero si sposta una leva di tanti scatti quanti ne indica ciascuna cifra. Per l'addizione si imposta un addendo e, con un giro di manovella, lo si trasferisce al totalizzatore, dove viene sommato automaticamente al numero già presente. Girando la manovella in senso opposto si esegue la sottrazione. Per moltiplicare si imposta il moltiplicando e lo si somma ripetutamente secondo le cifre del moltiplicatore. Per esempio, moltiplicare per 45 significa eseguire cinque addizioni del moltiplicando e poi quattro addizioni dello stesso numero spostato di una posizione decimale.

Per dividere si colloca il dividendo nel totalizzatore e si sottrae ripetutamente il divisore, procedendo posizione per posizione. Il numero di giri di manovella effettuati in ciascuna posizione determina la corrispondente cifra del quoziente; l'eventuale residuo dà il resto.

Questa breve descrizione serve a comprendere la direzione dei successivi perfezionamenti. Alcuni aumentarono soltanto la comodità d'uso, senza modificare lo schema fondamentale. L'introduzione dell'azionamento elettrico, per esempio, accelerò il funzionamento e liberò l'operatore dal compito di girare la manovella.

Per rendere più rapida e semplice l'impostazione dei numeri vennero introdotte tastiere: invece di spostare una leva di un determinato numero di scatti, bastava premere il tasto della cifra desiderata. Si costruirono calcolatrici nelle quali l'operatore doveva soltanto introdurre i numeri e premere il tasto dell'operazione; la macchina proseguiva da sola. La velocità aumentò notevolmente: nei modelli più recenti all'epoca dell'opera, una moltiplicazione richiedeva circa un secondo dopo l'impostazione. Un'ulteriore accelerazione del solo meccanismo avrebbe dato poco vantaggio, perché l'operatore impiegava assai più tempo a premere i tasti e a trascrivere i risultati.

Macchine a schede perforate e macchine a relè.

Le macchine a schede perforate furono inventate per calcoli statistici e per usi finanziari e industriali. Erano destinate a eseguire grandi quantità di calcoli semplici e ripetitivi. La piccola memoria operativa e le limitate possibilità di programmazione le rendevano meno comode per i calcoli scientifici e tecnici. Nonostante questi limiti, prima delle macchine elettroniche veloci furono largamente impiegate anche in calcoli complessi e di grandi dimensioni, quando il procedimento poteva essere ridotto a una breve sequenza di operazioni da ripetere molte volte, per esempio nella costruzione di tavole.

I numeri sono registrati su schede perforate, come nella figura 2. Cifre e simboli sono rappresentati da fori in posizioni prestabilite. La scheda passa attraverso un sistema di spazzole di contatto: quando una spazzola incontra un foro, chiude un circuito elettrico che attiva una determinata operazione della macchina.

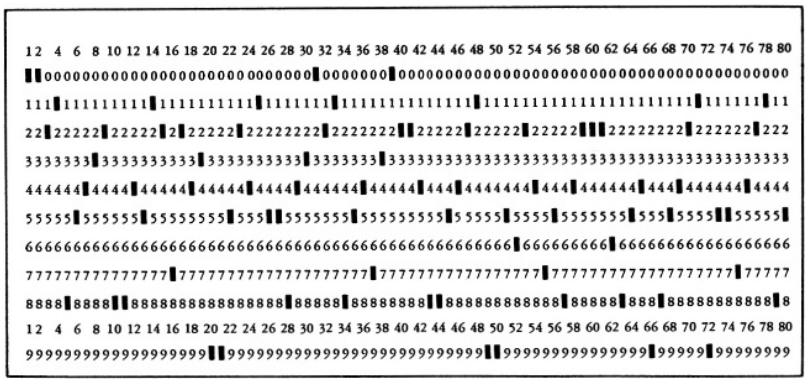


Figura 2.

I vari tipi di macchine sono progettati per lavorare in complessi che comprendono almeno gli apparecchi seguenti. Le velocità riportate si riferiscono alle macchine descritte nell'epoca del testo.

La perforatrice serve a praticare i fori nelle schede. È azionata manualmente mediante una tastiera e lavora a una velocità paragonabile a quella di una macchina per scrivere.

La selezionatrice dispone le schede nell'ordine in cui devono essere introdotte nelle macchine di calcolo. La velocità è di 450-650 schede al minuto.

La riproduttrice trasferisce perforazioni da una scheda a un'altra, confronta due gruppi di schede e seleziona quelle con determinate perforazioni. La velocità è di circa 100 schede al minuto.

La tabulatrice esegue addizioni e sottrazioni e stampa i risultati. Può elaborare da 6.000 a 9.000 schede all'ora.

La perforatrice moltiplicatrice esegue addizioni, sottrazioni e moltiplicazioni, registrando i risultati mediante perforazioni. Con numeri di sei o sette cifre può eseguire 700-1.000 moltiplicazioni all'ora.

Queste macchine lavorano piuttosto lentamente. Come valutazione approssimativa, il complesso appena descritto può sostituire da dodici a diciotto calcolatrici da tavolo. I primi tentativi di aumentare la velocità portarono alle macchine a relè elettromeccanici, circa dieci volte più rapide delle semplici macchine a schede. I vantaggi sotto altri aspetti furono notevoli: potevano eseguire programmi complessi e disponevano di un sistema di comando flessibile, che ampliò molto il campo dei problemi scientifici e tecnici affrontabili. La loro comparsa coincise però quasi con quella dei primi calcolatori elettronici a comando programmato, che produssero un ulteriore forte aumento della velocità. Per dare un'idea delle possibilità aperte dall'elettronica, il tempo di commutazione di una valvola elettronica si misura in milionesimi di secondo.

Macchine matematiche ad azione continua, o analogiche.

Queste macchine sono sistemi fisici - dispositivi meccanici, circuiti elettrici e così via - costruiti in modo che fra i loro parametri, variabili con continuità, valgano le stesse relazioni quantitative del problema matematico da risolvere. I parametri possono essere spostamenti, angoli di rotazione, correnti, tensioni e altre grandezze. Si parla perciò di macchine analogiche o di simulazione.

Ogni macchina ad azione continua è progettata per risolvere una classe relativamente ristretta di problemi.

La precisione dipende dalla qualità dei componenti, dal montaggio, dalla taratura, dagli errori dovuti all'inerzia e da altri fattori. L'esperienza sulle macchine descritte dall'autore indicava in genere una precisione di due o tre cifre significative. Sotto questo aspetto esse erano inferiori alle macchine numeriche: per queste ultime il numero di cifre può, in linea di principio, essere aumentato, purché si disponga delle risorse necessarie.

Una caratteristica importante delle macchine analogiche è la possibilità di risolvere numerosi problemi dello stesso tipo. In molti casi esse producevano la soluzione più rapidamente delle macchine numeriche allora disponibili. Il vantaggio principale stava spesso nella maggiore comodità di introdurre i dati iniziali e di ottenere il risultato in una forma adatta all'uso.

Esistono numerosi tipi di macchine analogiche. Si possono costruire dispositivi che riproducono operazioni matematiche quali addizione, moltiplicazione, integrazione e derivazione, oppure formule di calcolo: per esempio il valore di un polinomio o i coefficienti di Fourier nell'analisi armonica. Si possono anche simulare equazioni numeriche o funzionali. Problemi di rami scientifici molto diversi conducono spesso alle stesse equazioni differenziali. Questa identità permette, per esempio, di simulare fenomeni termici con sistemi elettrici e di risolvere problemi di termotecnica mediante misure elettriche, che possono risultare più precise e più semplici delle corrispondenti misure termiche.

La varietà di macchine analogiche è troppo grande per descriverle, o anche soltanto illustrarne tutti i principi costruttivi, in poche parole. Per dare almeno un'idea della simulazione di problemi matematici, descriviamo due semplici dispositivi: uno per integrare funzioni e uno per risolvere approssimativamente l'equazione di Laplace.

L'integratore ad attrito della figura 3 serve, come indica il nome, a integrare funzioni. Il principio è illustrato nella figura 4. Il componente 1 è il basamento; il 2 è un disco orizzontale con albero verticale; il 3 è una rotella a bordo liscio, che può rotolare sul disco e spostarsi radialmente, nel piano perpendicolare a quello di rotolamento. I componenti 4 e 5 costituiscono un meccanismo a vite: la vite 4 è collegata al carrello che porta la rotella. Indichiamo con h l'avanzamento della vite per radiante, cioè il passo ordinario diviso per 2π. Una rotazione di angolo γ sposta allora la rotella di una distanza radiale $\rho = h\gamma$ nel piano del disegno.

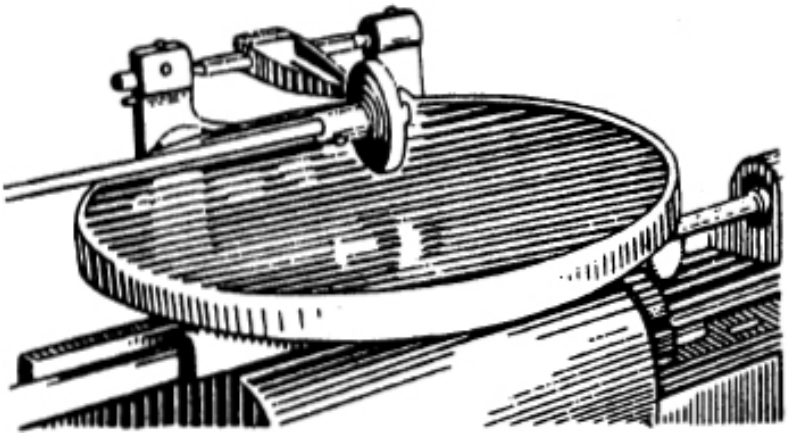


Figura 3

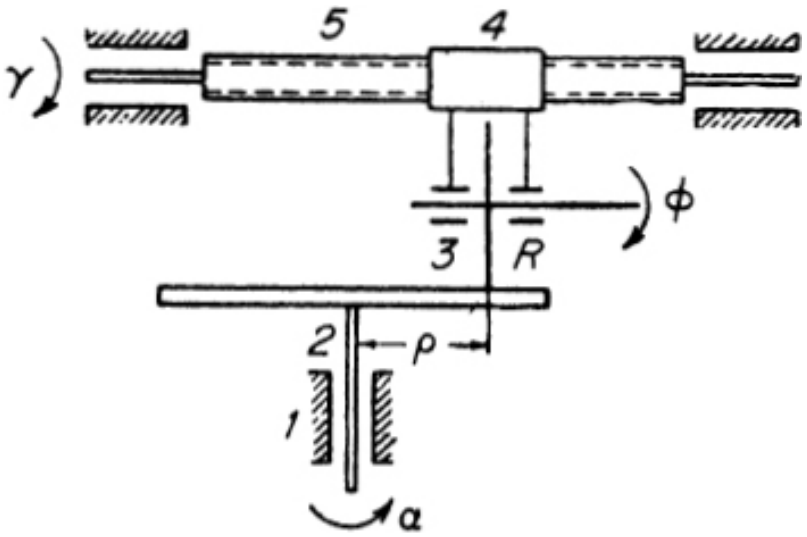


Figura 4

Se l'albero del disco ruota di un angolo $d\alpha$, il punto di contatto percorre un arco di lunghezza $\rho\, d\alpha$. Se la rotella, di raggio R, rotola senza slittare, il suo angolo di rotazione soddisfa

$$d\phi = \frac{\rho}{R}\, d\alpha = \frac{h}{R}\, \gamma\, d\alpha.$$

Supponiamo che la rotazione del disco inizi dall'angolo α_0 e che l'angolo iniziale della rotella sia ϕ_0. Integrando l'uguaglianza precedente si ottiene

$$\phi - \phi_0 = \frac{h}{R} \int_{\alpha_0}^{\alpha} \gamma\, d\alpha.$$

Scegliendo opportunamente la relazione fra γ e α, l'integratore permette di calcolare l'integrale desiderato in molti casi diversi. Combinando meccanismi integratori si può ottenere una soluzione meccanica di numerose equazioni differenziali.

Passiamo al secondo esempio. Sia dato nel piano un dominio Ω, delimitato da una curva l. Si cerca una funzione u che soddisfi all'interno l'equazione di Laplace

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

e assuma sul contorno i valori assegnati

$$u\,|_l = f.$$

Introduciamo una griglia quadrata di punti

$$x_j = x_0 + jh,\ y_k = y_0 + kh,$$
$$j, k = 0, \pm 1, \pm 2, \ldots .$$

e approssimiamo Ω con un poligono composto da quadrati della griglia. Alla curva l corrisponde una spezzata, sulla quale trasferiamo opportunamente i valori al contorno di f. Indichiamo con u_{jk} il valore incognito nel nodo (x_j, y_k). Per ottenere una soluzione approssimata sostituiamo l'equazione di Laplace con il sistema algebrico che, in ogni nodo interno, impone

$$u_{jk} = \tfrac{1}{4}(u_{j+1,k} + u_{j,k+1} + u_{j-1,k} + u_{j,k-1}).$$

Questo sistema si può risolvere costruendo un modello elettrico. Disponiamo nel piano una rete conduttrice bidimensionale come quella della figura 5, con la stessa resistenza fra ogni coppia di nodi adiacenti. Applichiamo ai nodi di bordo tensioni proporzionali ai valori assegnati di u, scegliendo le unità in modo che i valori numerici coincidano. Esse determinano le tensioni nei nodi interni. Indicando con $U_{j,k}$ la tensione del nodo (x_j, y_k), la legge di Kirchhoff dà

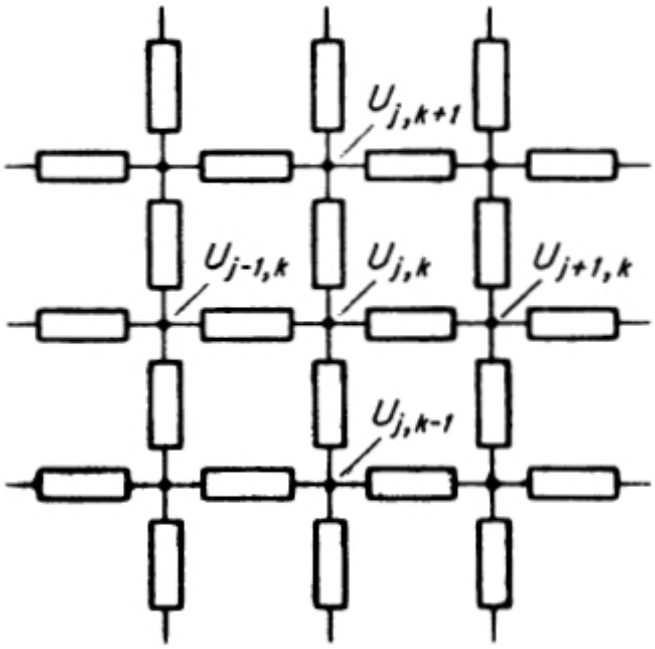


Figura 5.

$$\frac{1}{R}\left[(U_{j+1,k} - U_{j,k}) + (U_{j,k+1} - U_{j,k}) + (U_{j-1,k} - U_{j,k}) + (U_{j,k-1} - U_{j,k})\right] = 0,$$

che differisce dal sistema precedente soltanto nella notazione. I valori u_{jk} della soluzione discreta coincidono quindi, nella scala scelta, con le tensioni U_{jk}, misurabili sul modello con i consueti strumenti elettrici.

Letture consigliate

V. N. Faddeeva, *Computational methods of linear algebra*, Dover, New York, 1959.

D. R. Hartree, *Numerical analysis*, 2ª ed., Oxford University Press, New York, 1958.

F. B. Hildebrand, *Introduction to numerical analysis*, McGraw-Hill, New York, 1956.

A. S. Householder, *Principles of numerical analysis*, McGraw-Hill, New York, 1953.

W. J. Karplus e W. W. Soroka, *Analog methods: computation and simulation*, 2ª ed., McGraw-Hill, New York, 1959.

W. E. Milne, *Numerical calculus: approximations, interpolation, finite differences, numerical integration and curve fitting*, Princeton University Press, Princeton, N. J., 1949.

W. E. Milne, *Numerical solution of differential equations*, Wiley, New York, 1953.

A. M. Ostrowski, *Solution of equations and systems of equations*, Academic Press, New York, 1960.

J. B. Scarborough, *Numerical mathematical analysis*, 5ª ed., Johns Hopkins Press, Baltimore, Md., 1962.

R. G. Stanton, *Numerical methods for science and engineering*, Prentice-Hall, Englewood Cliffs, N. J., 1961.

Note al capitolo XIII

[1] Si veda il capitolo XII, §3.

[2] Ricordiamo due semplici esempi di procedimenti lentamente convergenti. È noto che la serie

$$\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

converge al logaritmo naturale di 2. Possiamo dunque calcolare approssimativamente ln 2 mediante la somma

$$s_n = \frac{1}{1} - \frac{1}{2} + \cdots \pm \frac{1}{n}$$

dei primi n termini, scegliendo n abbastanza grande. Per ottenere un errore inferiore a metà dell'unità nella quinta cifra significativa occorrono circa 100.000 termini. Sommarne così tanti con una calcolatrice da tavolo sarebbe molto laborioso. Un altro esempio noto è la serie

$$\frac{1}{\sqrt{2}} = 1 - \frac{1}{2 \cdot 1!} + \frac{1 \cdot 3}{2^2 \cdot 2!} - \frac{1 \cdot 3 \cdot 5}{2^4 \cdot 3!} + \frac{1 \cdot 3 \cdot 5 \cdot 7}{2^4 \cdot 4!} - \cdots.$$

La convergenza è così lenta che, per calcolare $1/\sqrt{2}$ con precisione dell'ordine di 10^{-5}, il numero di termini richiesto è dell'ordine di 10^{10}. Una mole simile di lavoro era difficile anche per le macchine veloci dell'epoca.

[3] Per la loro interpretazione geometrica, questo teorema e altri analoghi vengono spesso chiamati teoremi delle contrazioni.

[4] In questo paragrafo descriviamo soltanto gli strumenti e le macchine più semplici. I calcolatori veloci contemporanei alla stesura dell'opera sono trattati nel capitolo XIV. Per ragioni di spazio omettiamo anche i metodi grafici.

[5] Una colonna può essere molto lunga e venire suddivisa in più colonne per comodità di stampa; la tavola resta naturalmente a una sola entrata.

[6] Di regola l'interpolazione è più complessa quando gli argomenti tabulati x_i sono distanti, più semplice quando sono vicini. Anche la rapidità richiesta può variare molto. Nelle tavole storicamente destinate all'artiglieria, l'interpolazione doveva essere quasi immediata, «a vista». Nelle tavole scientifiche di maggiore precisione si poteva invece ammettere una successione di operazioni.

Nota editoriale

La traduzione conserva le due sezioni, le cinque figure, le quattro tavole numerate, le formule, la bibliografia e le sei note. Le tavole 1 e 3 sono ricomposte con intestazioni italiane e virgola decimale, mantenendo i valori dell'originale. Le altre tavole e gli schemi matematici sono riprodotti. Il richiamo alla nota 4 è spostato dal titolo del §2 all'inizio del testo per renderlo direttamente accessibile.

Nel §1 il polinomio di Taylor di grado cinque per sin x ha, su $[0, \pi/4]$, errore non superiore a $(\pi/4)^7/7!$, circa $3{,}66 \cdot 10^{-5}$: la precisione 10^{-5} indicata nell'originale è sostituita con $4 \cdot 10^{-5}$. Per ottenere il limite originario si potrebbe invece conservare anche il termine $-x^7/7!$. Nella (7) l'argomento del nucleo è x_j, non y_j; il rinvio alla rappresentazione mediante coefficienti rimanda alla (5). Sono esplicitate le ipotesi di convergenza e adeguatezza del sistema di funzioni nel procedimento di Galerkin.

Nella dimostrazione del teorema delle contrazioni si ripristina il valore assoluto di $x_3 - x_2$ e si corregge la parentesi della stima $m\,q^n/(1 - q)$. L'interpretazione geometrica riguarda le distanze fra immagini di coppie di punti, senza supporre la monotonia della trasformazione. Il passaggio da $x = (1/2)\tan x$ a $x = \arctan(2x)$ è limitato al ramo della radice cercata; si indica un intervallo invariante. Il valore 1,16556 è approssimato e la ripetizione delle cifre nelle iterazioni è riferita alla precisione usata. La soluzione dell'equazione $y' = y$ è scritta correttamente $y = e^x$. Le tavole illustrano separatamente gli effetti dell'iterazione, della discretizzazione e degli arrotondamenti.

Nel confronto fra determinanti ed eliminazione, si precisa $\Delta \neq 0$ e si corregge l'indice della colonna sostituita in Δ_j. Per lo schema descritto, conservando i coefficienti normalizzati, l'eliminazione richiede $\sum_{k=1}^{n} k^2$ moltiplicazioni e divisioni, e la sostituzione all'indietro $n(n - 1)/2$ moltiplicazioni. Il totale è $n(2n^2 + 6n - 2)/6$: 65 per $n = 5$ e 430 per $n = 10$. Questi conteggi sostituiscono 75 e 475 dell'originale; non comprendono addizioni e sottrazioni. La formula di Cramer resta valida, ma il confronto riguarda il calcolo diretto dei determinanti mediante tutti i termini della definizione. Si ammettono scambi di righe per evitare divisori nulli.

Nel §2 si precisa l'ipotesi di invertibilità necessaria a combinare le tavole, e si scrive coerentemente la funzione composta $f(\phi(x,y),z)$. Nel modello meccanico h indica l'avanzamento della vite per radiante: con il passo ordinario p si avrebbe $h = p/(2\pi)$. È ripristinato l'angolo iniziale α_0 e sono distinti gli indici indipendenti della griglia quadrata. «Macchine a schede perforate» rende il termine storico inglese usato per quella specifica categoria, senza identificarla con tutti i calcolatori digitali. Le prestazioni, i confronti fra tecnologie e le valutazioni di fattibilità descrivono l'epoca dell'opera. Sono corretti refusi nei nomi Morland, Leibniz e Odhner.

CAPITOLO XIV

Le macchine calcolatrici elettroniche

S. A. Lebedev e L. V. Kantorovič

§1. Scopi e principi fondamentali dei calcolatori elettronici

I metodi matematici sono largamente usati nella scienza e nella tecnica, ma molti problemi importanti richiedono una mole di calcoli tale da risultare praticamente insolubili con una comune calcolatrice da tavolo. La comparsa delle macchine elettroniche, capaci di una rapidità prima sconosciuta, ha rivoluzionato l'applicazione della matematica ai principali problemi della fisica, della meccanica, dell'astronomia, della chimica e di altre discipline.

Un calcolatore elettronico universale contemporaneo alla stesura di questo capitolo eseguiva migliaia, e persino decine di migliaia, di operazioni aritmetiche e logiche al secondo, sostituendo il lavoro di centinaia di migliaia di calcolatori umani. Una simile velocità permetteva, per esempio, di calcolare la traiettoria di un missile in un tempo inferiore a quello del suo volo.

Oltre alla rapidità, i calcolatori universali offrono la possibilità di risolvere problemi molto diversi con una stessa macchina. Essi costituiscono uno strumento qualitativamente nuovo: non aumentano soltanto la produzione di risultati già ottenibili, ma rendono affrontabili problemi prima considerati inaccessibili.

In molti casi i risultati hanno valore soltanto se vengono ottenuti rapidamente. È evidente, per esempio, nelle previsioni meteorologiche per il giorno successivo. Con calcolatrici manuali, i calcoli necessari a una previsione attendibile possono richiedere essi stessi parecchi giorni, perdendo così ogni utilità pratica. I calcolatori elettronici permettono invece di ottenere i risultati in tempo utile.

La macchina elettronica ad alta velocità.

Un esempio è il calcolatore BESM, costruito presso l'Istituto di meccanica di precisione e tecnica del calcolo dell'Accademia delle scienze dell'URSS. La macchina eseguiva fra 8.000 e 10.000 operazioni aritmetiche al secondo. Un operatore esperto, con una calcolatrice da tavolo, poteva eseguirne circa 2.000 in una giornata lavorativa. In poche ore, quindi, il calcolatore elettronico svolgeva un lavoro che l'operatore non avrebbe terminato in tutta la vita. Una sola macchina sostituiva un esercito di decine di migliaia di operatori, che avrebbero richiesto una superficie di centinaia di migliaia di iarde quadrate anche soltanto per essere ospitati.

Queste macchine furono impiegate per numerosi problemi scientifici e tecnici. Secondo le valutazioni riportate nell'opera, ne derivarono risparmi per centinaia di milioni di dollari. Diamo alcuni esempi.

Per l'effemeride astronomica internazionale furono calcolate in pochi giorni le orbite di circa settecento asteroidi, tenendo conto delle perturbazioni di Giove e Saturno. Si determinarono le coordinate per dieci anni, riportando le posizioni a intervalli di quaranta giorni. In precedenza un ufficio di calcolo avrebbe impiegato molti mesi.

La costruzione di carte sulla base di rilievi geodetici richiede sistemi algebrici con numerose incognite. Problemi con 800 equazioni, che comportavano fino a 250 milioni di operazioni aritmetiche, furono risolti dalla macchina in meno di venti ore.

Sullo stesso calcolatore furono costruite tavole della massima pendenza compatibile con la stabilità delle sponde di un canale, ottenendo notevoli risparmi di tempo e materiali nella costruzione di centrali idroelettriche. In precedenza quindici calcolatori umani avevano lavorato per mesi senza riuscire a risolvere neppure un caso particolare. La macchina completò i calcoli per dieci casi in meno di tre ore.

Una macchina permette di esaminare rapidamente molte soluzioni e scegliere la più adatta: per esempio la struttura di un ponte, la forma di un'ala, di un ugello per motore a reazione o di una pala di turbina.

La possibilità di aumentare molto la precisione dei calcoli consente di costruire rapidamente tavole per la scienza e la tecnica. Sul BESM, una tavola di 50.000 valori dell'integrale di Fresnel richiese soltanto un'ora.

Applicazioni ai problemi logici.

Oltre ai problemi matematici, un calcolatore può affrontare problemi logici: per esempio tradurre testi da una lingua all'altra. In questo caso si memorizzano codifiche di parole e informazioni che svolgono il ruolo di un dizionario.

Confrontando le parole del testo con quelle del dizionario, la macchina trova i corrispondenti nella lingua desiderata. Applica poi regole grammaticali e sintattiche descritte nel programma: modifica caso, numero o tempo verbale e dispone le parole nell'ordine richiesto dalla frase. Il testo tradotto viene stampato. Per preparare programmi capaci di buone traduzioni occorre un vasto e paziente lavoro di linguisti e matematici.

Presso l'Accademia delle scienze dell'URSS furono preparati dizionari e programmi sperimentali per tradurre testi scientifici e tecnici dall'inglese al russo. Alla fine del 1955 il BESM produsse una prima traduzione sperimentale, benché non fosse una macchina specificamente progettata per questo compito.

Sul BESM furono affrontati sperimentalmente anche problemi logici complessi, per esempio problemi di scacchi. L'enorme numero di combinazioni escludeva un'analisi completa del gioco con le macchine dell'epoca. Si ricorreva quindi a valutazioni approssimate: per esempio 10.000 punti per il re, 100 per la donna, 50 per una torre, oltre a punteggi per vantaggi

posizionali come colonne aperte o pedoni passati. Esaminando diverse continuazioni, la macchina sceglieva quella che, dopo un certo numero di mosse e tenendo conto delle risposte dell'avversario, garantiva il punteggio migliore. Il numero enorme di combinazioni limitava però la ricerca a poche mosse, impedendo di studiare piani strategici di lungo periodo.

Principi fondamentali di funzionamento.

I calcolatori descritti in questo capitolo sono complessi sistemi di componenti elettronici: valvole, elementi semiconduttori al germanio, elementi magnetici, fotocellule, resistori, condensatori e altri dispositivi della radiotecnica.

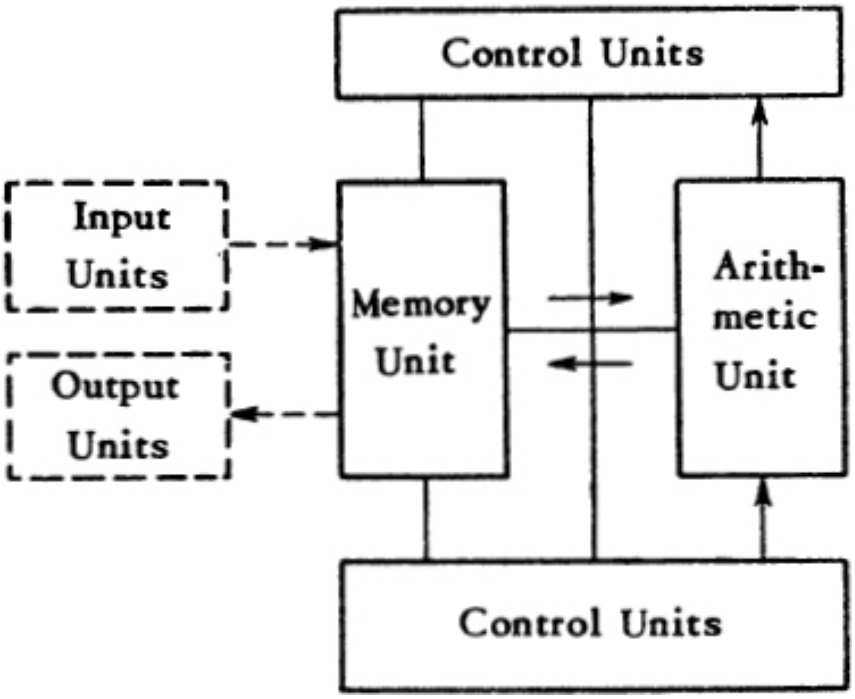


Figura 1. Schema delle unità fondamentali di un calcolatore elettronico. Input Units: unità di ingresso; Output Units: unità di uscita; Memory Unit: memoria; Arithmetic Unit: unità aritmetica; Control Units: unità di controllo.

Le operazioni aritmetiche sono eseguite con grande rapidità da circuiti riuniti nell'unità aritmetica (figura 1).

Ma la sola rapidità delle operazioni non basta. L'intero processo deve essere automatico: anche il reperimento dei numeri e la scelta della sequenza di operazioni devono avvenire senza intervento manuale.

I numeri da elaborare e i risultati intermedi devono essere conservati nella macchina. A questo serve l'unità di memoria, che permette di recuperare i valori richiesti e registrare i risultati. La sua capacità, cioè il numero di valori che può contenere, determina in larga misura la versatilità del calcolatore.

Nei calcolatori considerati qui, la memoria conteneva generalmente da 1.000 a 4.000 numeri.

L'unità di controllo dirige il prelievo dei numeri dalla memoria, l'operazione da eseguire, la registrazione del risultato e il passaggio all'operazione successiva. Dopo l'introduzione del programma e dei dati iniziali, essa rende automatico l'intero procedimento.

Apposite unità di ingresso e di uscita servono a introdurre dati e programmi e a stampare i risultati.

Per usare la macchina bisogna poter verificare l'attendibilità dei risultati. Il controllo può essere affidato a dispositivi appositi oppure a verifiche logiche e matematiche inserite in un programma. L'esempio più semplice è il controllo per duplicazione, tradizionalmente detto calcolo «a due mani»: si ripete il calcolo e si confrontano i risultati.

Prima di risolvere un problema bisogna formularlo matematicamente a partire dal fenomeno fisico: mediante formule algebriche, equazioni differenziali o integrali, oppure altre relazioni. I metodi dell'analisi numerica permettono poi, in moltissimi casi, di ricondurlo a una sequenza di operazioni aritmetiche. Anche problemi assai complessi vengono così affrontati mediante le quattro operazioni.

Per eseguire manualmente un'operazione aritmetica, prendiamo due numeri, compiamo l'operazione e scriviamo il risultato, che può essere una risposta finale oppure un dato per i calcoli successivi.

Lo stesso accade nel calcolatore. La memoria è composta da celle numerate; per selezionare un numero occorre indicare la cella in cui è conservato.

Per un'operazione su due numeri si devono specificare le celle da cui prelevarli, l'operazione da eseguire e la cella in cui collocare il risultato. Queste informazioni, espresse in un codice prestabilito, costituiscono un'istruzione.

Risolvere un problema significa eseguire una successione di istruzioni. Esse costituiscono il programma e, di solito, vengono anch'esse conservate nella memoria.

Il programma, cioè l'insieme delle istruzioni che realizza la sequenza di operazioni necessaria, viene preparato in anticipo dai matematici.

Molti problemi richiedono centinaia di milioni di operazioni. Si usano quindi procedimenti che permettono a un numero relativamente piccolo di istruzioni di governare una grande mole di calcolo.

Accanto alle istruzioni aritmetiche esistono istruzioni logiche: per esempio il confronto fra due numeri, che permette di scegliere fra due prosecuzioni del calcolo a seconda di quale sia maggiore.

Programma e dati iniziali sono scritti secondo un codice concordato. Nei sistemi descritti, le istruzioni sono registrate come perforazioni su schede o nastri, oppure come segnali su nastro magnetico. I codici vengono introdotti nella memoria e la macchina esegue automaticamente il programma.

Anche i risultati possono essere registrati su nastro magnetico. Dispositivi di decodifica e stampa li trasformano in cifre ordinarie e li stampano in forma di tavola.

La rapidità dei calcolatori procura un risparmio di lavoro intellettuale paragonabile a quello di lavoro fisico prodotto dalle macchine industriali. Gli autori sottolineano tuttavia che il calcolatore esegue un programma predisposto dall'uomo e, nella concezione qui esposta, non possiede di per sé capacità creative né sostituisce l'essere umano.

La diffusione dei calcolatori negli istituti scientifici e tecnici, negli uffici di progettazione e negli organismi di pianificazione apre prospettive vastissime per l'economia. Ingegneri e matematici trovano nuovi campi di attività nello sviluppo delle macchine, nel loro funzionamento e nelle loro applicazioni.

Le macchine elettroniche sono strumenti potenti nelle mani dell'uomo; gli autori ne sottolineano la grande importanza per l'economia nazionale.

§2. Programmazione e codifica per i calcolatori elettronici veloci

Principi della programmazione: 1) Il metodo di Eulero.

Il metodo matematico scelto per approssimare la soluzione viene tradotto in una successione di operazioni aritmetiche, eseguita dalla macchina mediante le istruzioni del programma. Se occorresse un'istruzione distinta per ogni operazione, il programma sarebbe lunghissimo e scriverlo richiederebbe quasi quanto eseguire i calcoli a mano. Bisogna dunque fare in modo che poche istruzioni bastino per molte operazioni.

Per chiarire la struttura delle istruzioni e la preparazione dei programmi, esaminiamo anzitutto le operazioni necessarie a risolvere manualmente un problema molto semplice.

Consideriamo il metodo di Eulero per l'equazione differenziale del primo ordine con la condizione iniziale

$$\frac{dy}{dx} = ay, \qquad y|_{x_0} = y_0 . \tag{1}$$

Si divide l'intervallo della variabile x in parti uguali di lunghezza $\Delta x = h$. In ciascuna parte si considera la derivata costante e uguale al valore all'inizio della parte stessa.[1] Il calcolo relativo al passo k è allora

$$\left(\frac{dy}{dx}\right)_k = ay_k\,,$$

$$\Delta y_k = \left(\frac{dy}{dx}\right)_k h = (ah)y_k\,,$$

$$y_{k+1} = y_k + \Delta y_k\,,$$

$$x_{k+1} = x_k + h.$$

Completato il passo *k*, si passa al successivo. Il calcolo comincia dai dati x_0, y_0. La sequenza delle operazioni è riportata nella tavola 1.

Nel calcolo manuale si scrivono soltanto le prime tre operazioni; le altre, come ripetere il procedimento per l'intervallo successivo o terminare il calcolo, restano sottintese. Per la macchina anche queste devono essere formulate esplicitamente: sono le operazioni di controllo 4-7. Alcune hanno carattere incondizionato, come la stampa dei risultati; altre dipendono dal risultato raggiunto. Le operazioni 6 e 7 si escludono a vicenda e vengono riunite in un confronto: «Se *x* è minore di x_n, ripetere dalla prima operazione; se è maggiore o uguale a x_n, terminare». La prosecuzione del calcolo dipende così dal valore di *x* ottenuto.

Tavola 1. Operazioni per risolvere la (1) con il metodo di Eulero

N.	Grandezza	Formula	Calcoli
1	Δy_k	$(ah)y_k$	$(ah)(2)_{k-1}$
2	y_{k+1}	$y_k + \Delta y_k$	$(2)_{k-1} + (1)_k$
3	x_{k+1}	$x_k + h$	$(3)_{k-1} + h$
4			Stampare il valore di x_{k+1}.
5			Stampare il valore di y_{k+1}.
6			Se $x < x_n$, ripetere dall'operazione 1 con i nuovi valori di x e y.
7			Quando $x \geq x_n$, terminare.

Per la notazione dei calcoli si veda la nota [2].

2. Il sistema a tre indirizzi.

La tavola 1 mostra che ogni operazione richiede di indicare quale operazione eseguire, su quali numeri e dove collocare il risultato, che potrà essere usato nei calcoli successivi.

Poiché i numeri sono codificati e conservati nella memoria, bisogna indicare gli indici delle celle da cui prelevarli e della cella di destinazione. Nasce così il sistema naturale di istruzioni a tre indirizzi.

Nel codice di ogni istruzione un gruppo di posizioni specifica l'operazione: è il codice operativo. Le posizioni rimanenti si dividono in tre gruppi uguali, gli indirizzi (figura 2). Il

primo indica la cella del primo operando, il secondo quella del secondo operando, il terzo la cella in cui depositare il risultato.

Codice operativo	1° indirizzo	2° indirizzo	3° indirizzo

Figura 2. Struttura di un'istruzione a tre indirizzi.

Anche le istruzioni di controllo possono essere rappresentate così. Per «trasferire un numero alla stampa», il codice operativo identifica la stampa, il primo indirizzo indica la cella del numero e il terzo l'unità di stampa; il secondo indirizzo resta inutilizzato. L'istruzione che sceglie fra due prosecuzioni si chiama confronto. Essa confronta i numeri indicati dai primi due indirizzi: se il primo è minore del secondo, si passa all'istruzione indicata dal terzo indirizzo; altrimenti si prosegue con l'istruzione immediatamente successiva.

Le istruzioni, come i numeri, sono conservate in memoria e vengono eseguite nell'ordine delle celle, salvo un cambiamento di percorso disposto, per esempio, da un confronto.

Vediamo il programma dell'esempio precedente. Disponiamo i dati nella memoria così:

La quantità ah nella cella 11.

La quantità h nella cella 12.

Il valore finale x_n nella cella 13.

Il valore corrente di x nella cella 14.

Il valore corrente di y nella cella 15.

La cella di lavoro[3] è la 16.

Alla tavola precedente corrisponde il programma della tavola 2.

Tavola 2. Programma per il metodo di Eulero

N.	Operazione	1° ind.	2° ind.	3° ind.	Significato
1	Moltiplica	11	15	16	$\Delta y_k = (ah)y_k$
2	Somma	15	16	15	$y_{k+1} = y_k + \Delta y_k$
3	Somma	14	12	14	$x_{k+1} = x_k + h$
4	Stampa	14	-	1	Stampare x_{k+1} nella prima unità di stampa.
5	Stampa	15	-	2	Stampare y_{k+1} nella seconda unità di stampa.
6	Confronta	14	13	1	Se $x < x_n$, tornare alla 1; altrimenti passare alla 7.
7	Arresta	-	-	-	Fine del calcolo.

Le istruzioni sono collocate nelle celle 1-7. L'unità di controllo riceve inizialmente quella della cella 1: moltiplica il numero nella cella 11 per quello nella 15, calcolando $\Delta y_k = (ah)y_k$, e deposita il risultato nella cella di lavoro 16. Passa poi alla seconda istruzione, che calcola y_{k+1}

= $y_k + \Delta y_k$ e lo registra nella cella 15, sostituendo il precedente valore di y. La terza aggiorna x; la quarta e la quinta stampano i nuovi valori. La sesta confronta la cella 14 con la 13, cioè x_{k+1} con il valore finale x_n. Se il primo è minore, il calcolo deve continuare per un altro intervallo e ritorna all'istruzione 1, indicata nel terzo indirizzo del confronto. Se invece $x_{k+1} \geqslant x_n$, si passa all'istruzione 7, che arresta la macchina.

Prima di cominciare bisogna caricare le istruzioni nelle celle 1-7, le costanti nelle celle 11-13 e i dati iniziali x_0, y_0 nelle celle 14 e 15.

Dopo il primo intervallo, le celle 14 e 15 contengono x_1 e y_1: sono i dati iniziali per il passo seguente. Il medesimo programma può quindi essere ripetuto.

L'esempio mostra come la ripetizione ciclica di una sequenza di istruzioni produca molti calcoli con un programma relativamente breve. Questo procedimento è largamente utilizzato nella programmazione.

3. Modifica degli indirizzi.

Un secondo procedimento, molto usato nelle macchine descritte, riduce la lunghezza del programma modificando automaticamente gli indirizzi di alcune istruzioni. Lo illustriamo con il calcolo di un polinomio.

Si voglia calcolare il valore del polinomio

$$y = a_0x^6 + a_1x^5 + a_2x^4 + a_3x^3 + a_4x^2 + a_5x + a_6 .$$

Per il calcolo automatico conviene scriverlo nella forma annidata

$$y = (((((a_0x + a_1)x + a_2)x + a_3)x + a_4)x + a_5)x + a_6 .$$

Disponiamo i coefficienti a_0, ..., a_6 nelle celle 20-26 e x nella cella 31. Il programma, di facile costruzione, è riportato nella tavola 3.

Moltiplicazioni e addizioni si alternano. Tranne la prima, tutte le moltiplicazioni sono identiche: moltiplicare il contenuto della cella 27 per quello della 31 e lasciare il risultato nella 27. Anche le addizioni hanno il primo e il terzo indirizzo uguali; cambia soltanto il secondo, che cresce di uno ogni volta: cella 21 nella seconda istruzione, 22 nella quarta e così via.

Tavola 3. Calcolo di un polinomio

N.	Operazione	1° ind.	2° ind.	3° ind.	Risultato nella cella 27
1	Moltiplica	20	31	27	a_0x
2	Somma	27	21	27	$a_0x + a_1$
3	Moltiplica	27	31	27	$(a_0x + a_1)x$
4	Somma	27	22	27	$(a_0x + a_1)x + a_2$
5	Moltiplica	27	31	27	$((a_0x + a_1)x + a_2)x$
6	Somma	27	23	27	$((a_0x + a_1)x + a_2)x + a_3$
7	Moltiplica	27	31	27	$(((a_0x + a_1)x + a_2)x + a_3)x$
8	Somma	27	24	27	$(((a_0x + a_1)x + a_2)x + a_3)x + a_4$
9	Moltiplica	27	31	27	$((((a_0x + a_1)x + a_2)x + a_3)x + a_4)x$
10	Somma	27	25	27	$((((a_0x + a_1)x + a_2)x + a_3)x + a_4)x + a_5$
11	Moltiplica	27	31	27	$(((((a_0x + a_1)x + a_2)x + a_3)x + a_4)x + a_5)x$
12	Somma	27	26	27	$y = (((((a_0x + a_1)x + a_2)x + a_3)x + a_4)x + a_5)x + a_6$

Il programma si accorcia molto se l'indice nel secondo indirizzo dell'addizione viene modificato automaticamente. Le istruzioni sono conservate in memoria come codici e possono essere trattate anch'esse come numeri: aggiungendo un valore opportuno, se ne modifica un indirizzo. Si ottiene così il programma della tavola 4.

Tavola 4. Calcolo del polinomio con modifica degli indirizzi

N.	Operazione	1° indirizzo	2° indirizzo	3° indirizzo
1	Somma / copia	20	-	27
2	Moltiplica	27	31	27
3	Somma	27	21	27
4	Somma	3	28	3
5	Confronta	3	29	2
6	Arresta	-	-	-

La prima istruzione trasferisce il contenuto della cella 20 nella 27, così da rendere uniforme la moltiplicazione. La seconda e la terza producono $a_0x + a_1$. Per continuare bisogna aumentare di uno il secondo indirizzo della terza istruzione. Lo fa la quarta: prende il contenuto della cella 3, cioè il codice dell'addizione, vi aggiunge il contenuto della cella 28 e rimette il risultato nella cella 3. Per ottenere l'incremento desiderato, la cella 28 deve contenere

Codice operativo	1° indirizzo	2° indirizzo	3° indirizzo
-	-	1	-

Dopo questa operazione la terza istruzione assume la forma

Codice operativo	1° indirizzo	2° indirizzo	3° indirizzo
Somma	27	22	27

Il nuovo codice sostituisce quello precedente nella cella 3.

Possiamo ora ripetere il calcolo dalla moltiplicazione, cioè dall'istruzione 2. A questo serve il confronto dell'istruzione 5: esso confronta il nuovo contenuto della cella 3 con il valore di arresto conservato nella cella 29, che è

Codice operativo	1° indirizzo	2° indirizzo	3° indirizzo
Somma	27	27	27

All'inizio il contenuto della cella 3 è minore di quello della 29: il confronto rimanda quindi all'istruzione 2, indicata nel suo terzo indirizzo. Moltiplicazione e addizione si ripetono automaticamente; ogni volta la quarta istruzione aumenta di uno il secondo indirizzo dell'addizione.

Il ciclo continua finché tale indirizzo raggiunge 27, dopo sei esecuzioni. A quel punto la terza istruzione ha la forma

Codice operativo	1° indirizzo	2° indirizzo	3° indirizzo
Somma	27	27	27

e il suo codice coincide con quello nella cella 29. Il confronto rileva l'uguaglianza e passa all'istruzione successiva, la sesta, che termina il calcolo del polinomio.

La modifica automatica degli indirizzi, effettuata dal programma stesso, trovava largo impiego nei calcolatori descritti. Insieme ai cicli permetteva di eseguire una grande mole di operazioni con poche istruzioni.

4. Il sistema a un indirizzo.

Molte macchine usano, invece del sistema a tre indirizzi, istruzioni con un solo indirizzo oltre al codice operativo. Per operare su due numeri e registrare il risultato servono allora tre istruzioni: caricare il primo numero nell'unità aritmetica, introdurre il secondo ed eseguire l'operazione, infine trasferire il risultato in memoria. Spesso, però, il risultato serve soltanto come operando dell'operazione successiva: lo si può lasciare nell'unità aritmetica, evitando sia la registrazione sia il successivo caricamento. Nei programmi considerati dagli autori, il numero di istruzioni risultava perciò all'incirca doppio rispetto al sistema a tre indirizzi. Poiché ogni istruzione è più corta, l'occupazione di memoria era simile: una cella poteva spesso contenere due istruzioni a un indirizzo. Nel confrontare le velocità occorre tener conto di questa differenza; a parità di tempo per istruzione, il maggior numero di istruzioni poteva portare, negli esempi considerati, a tempi circa doppi. Si tratta di confronti fra quelle architetture, non di una regola generale.

Esistono anche sistemi di istruzioni a due o a quattro indirizzi.

5. Sottoprogrammi.

La soluzione di un problema si articola di solito in più fasi, molte delle quali ricorrono in problemi diversi: per esempio il calcolo di una funzione elementare per un argomento assegnato, o di un integrale definito di una funzione già calcolata.

Conviene preparare una volta per tutte sottoprogrammi standard per queste operazioni. Quando il programma principale ne ha bisogno, trasferisce l'esecuzione al sottoprogramma; terminato il compito, si ritorna al punto del programma principale dal quale si era partiti.

Una biblioteca di sottoprogrammi, registrati su schede o nastri magnetici, facilita molto il lavoro. Per numerosi problemi la programmazione si riduce così a brevi parti del programma principale che collegano opportunamente sottoprogrammi già disponibili.

6. Verifica dei risultati.

Un problema può richiedere milioni di operazioni, e un errore in una sola di esse può compromettere il risultato. Un controllo manuale completo è praticamente impossibile: deve intervenire la macchina stessa. Esistono dispositivi che verificano il funzionamento e arrestano il calcolo quando rilevano un errore; essi aumentano però la complessità dell'apparecchiatura e di solito non ne controllano ogni parte. Gli autori considerano particolarmente promettenti le verifiche inserite nel programma.

Il controllo più semplice consiste nella ripetizione indipendente del calcolo. La coincidenza dei risultati offre una verifica contro molti errori occasionali, ma non esclude errori sistematici ripetuti allo stesso modo. Per cercare questi ultimi si eseguono preventivamente calcoli di prova con risultati noti, coinvolgendo le diverse parti della macchina. Il buon esito delle prove aumenta la fiducia nel funzionamento, pur non costituendo una dimostrazione assoluta dell'assenza di qualsiasi errore.

Sono possibili controlli più elaborati, legati al problema. Per una traiettoria si possono calcolare separatamente le due componenti della velocità e il suo modulo, verificando a ogni passo la relazione

$$v^2 = v_x^2 + v_y^2 .$$

Nelle equazioni differenziali ordinarie si può ripetere il calcolo con passo $h/2$ oltre che con passo h. Il confronto aiuta a rilevare errori e a valutare l'adeguatezza del passo. In una tavola ottenuta per ricorrenza si possono ricalcolare alcuni valori di riferimento con metodi indipendenti; l'accordo fornisce un controllo sui valori intermedi, la cui affidabilità dipende anche dalla stabilità della ricorrenza. In altri casi si esaminano le differenze fra risultati successivi.

La preparazione del programma deve dunque comprendere una forma di verifica logica dei risultati.

Codifica di numeri e istruzioni.

Numeri e istruzioni vengono rappresentati nella macchina mediante codici. Nella maggior parte dei casi si usa il sistema binario anziché quello decimale.

Nel sistema decimale la base è 10. Ogni cifra assume uno dei valori da 0 a 9 e il peso di ogni posizione è dieci volte quello della precedente. Un intero non negativo si può quindi scrivere

$$N_{10} = k_0 10^0 + k_1 10^1 + k_2 10^2 + \cdots + k_n 10^n,$$

dove k_0, k_1, ..., k_n sono cifre da 0 a 9.

Nel sistema binario la base è 2, ogni cifra vale 0 oppure 1 e il peso di ogni posizione raddoppia. Si scrive dunque

$$N_2 = k_0 2^0 + k_1 2^1 + \cdots + k_p 2^p,$$

dove k_0, k_1, ..., k_p valgono 0 o 1.

I primi numeri si rappresentano nei due sistemi così:

Binario	0	1	10	11	100	101	110	111	1000	1001	1010	1011
Decimale	0	1	2	3	4	5	6	7	8	9	10	11

Per la parte frazionaria si usano analogamente potenze negative della base. Per esempio, $3\frac{1}{8}$ si scrive in binario

$$11.001.$$

La conversione fra sistemi di numerazione richiede operazioni aritmetiche che il calcolatore può eseguire mediante programmi appositi.

Le operazioni binarie seguono le stesse regole posizionali di quelle decimali. La somma di due unità dà zero nella posizione considerata e riporto di uno nella successiva. Per esempio,

$$1010 + 111 = 10001.$$

Moltiplicazione e divisione sono semplificate dal fatto che bastano le regole per moltiplicare per zero e per uno. Per esempio,

$$\begin{array}{r} 1010\ (10) \\ \times\ 101\ \ (5) \\ \hline 1010 \\ 0000 \\ 1010 \\ \hline 110010\ (50) \end{array}$$

Il sistema binario semplifica molto l'unità aritmetica, pur richiedendo più posizioni e, negli schemi elementari, più passi di moltiplicazione e divisione. Inoltre le sue cifre si rappresentano facilmente mediante due stati fisici: relè aperto o chiuso, presenza o assenza di segnale e così via.

Ogni cifra può corrispondere alla presenza o assenza di un segnale in un circuito dedicato, oppure allo stato di un relè: occorre allora un circuito per ciascuna cifra, come nella figura 3a. È la rappresentazione parallela. Oppure le cifre possono essere trasmesse a intervalli determinati lungo un solo circuito: è la rappresentazione seriale, mediante impulsi temporali. Gli intervalli sono scanditi da impulsi di sincronizzazione comuni alla macchina.

Nei sistemi paralleli tutte le cifre vengono trasmesse insieme, ciascuna sul proprio circuito. Nei sistemi seriali un solo circuito trasporta le cifre una dopo l'altra e il tempo di trasmissione cresce con il loro numero. A parità delle altre condizioni, il parallelismo aumenta la velocità ma richiede più apparecchiature.

Ogni macchina dispone di un numero determinato di posizioni per rappresentare i numeri. Occorre stabilire anche dove si trovi la virgola che separa parte intera e parte frazionaria; nel sistema binario si parla più precisamente di virgola binaria.

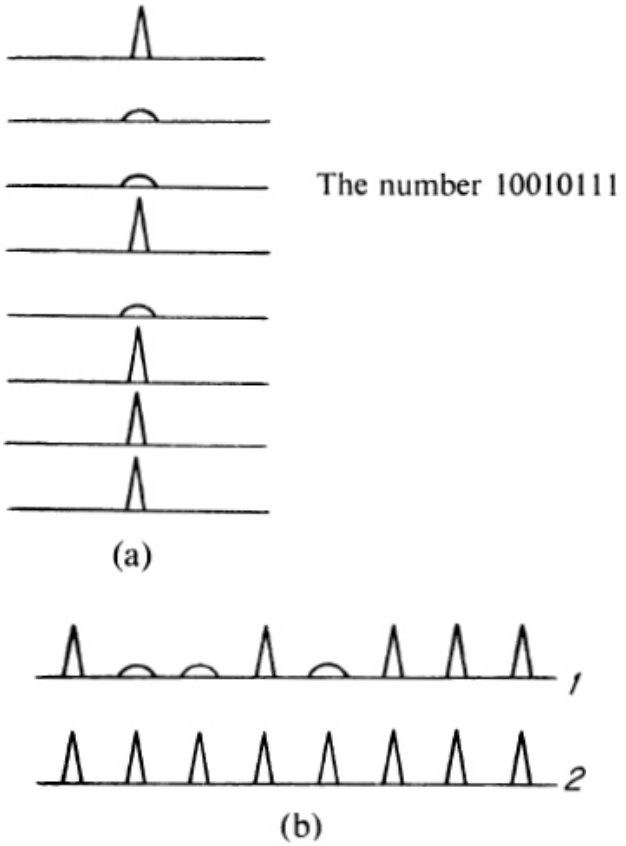


Figura 3. Rappresentazione dei codici: (a) parallela; (b) seriale. Il numero rappresentato è 10010111; in (b), la linea 1 mostra il codice e la linea 2 gli impulsi di sincronizzazione.

In alcune macchine la posizione della virgola è fissata: sono le macchine a virgola fissa. Nei modelli descritti si colloca spesso la virgola prima di tutte le cifre, imponendo numeri di valore assoluto minore di uno mediante un'opportuna scala. Nei calcoli complessi è difficile prevedere l'ampiezza dei risultati: scegliere una scala con molto margine riduce la precisione, mentre modificarla automaticamente durante il calcolo complica il programma.

Altre macchine indicano separatamente la posizione della virgola per ogni numero: sono le macchine a virgola mobile. Ciò equivale a rappresentare il numero mediante una parte significativa e un esponente:

$$N_{10} = 10^{k}N'_{10} \text{ nel sistema decimale;}$$
$$N_{2} = 2^{p}N'_{2} \text{ nel sistema binario.}$$

Per esempio, $97{,}35 = 10^{2} \cdot 0{,}9735$. Vanno quindi registrati sia l'esponente, p o k, sia le cifre significative. Nei limiti dell'intervallo rappresentabile, questo permette di utilizzare lo stesso numero di cifre significative per numeri di grandezza diversa, con precisione relativa dello stesso ordine. Si migliora così l'accuratezza, soprattutto nelle moltiplicazioni, e spesso non occorre scegliere esplicitamente una scala.

Questi vantaggi richiedono un'unità aritmetica più complessa, in particolare per addizione e sottrazione. Prima di sommare numeri con esponenti diversi bisogna allinearli; le ultime cifre del numero più piccolo possono essere scartate, come nell'esempio

$$10^2 \cdot 0{,}7587 + 10^0 \cdot 0{,}3743$$
$$\approx 10^2 \cdot 0{,}7587 + 10^2 \cdot 0{,}0037 = 10^2 \cdot 0{,}7624.$$

A virgola fissa, il codice binario coincide essenzialmente con la sequenza delle cifre del numero, qui di valore assoluto minore di uno:

$$.00110110000000 = \frac{27}{128}.$$

A virgola mobile, una parte del codice rappresenta l'esponente, anch'esso in binario. Per esempio,

$$6\frac{3}{4} = 2^3 \cdot \frac{27}{32} = 0011.11011000000.$$

Si riservano inoltre posizioni per il segno: per esempio 0 per il segno positivo e 1 per il negativo. Nella convenzione qui descritta vi è una posizione per il segno dell'esponente e una per il segno del numero.

Le istruzioni sono codificate nello stesso modo: alcuni bit indicano l'operazione e altri gli indirizzi delle celle di memoria.

§3. Principi tecnici delle unità di un calcolatore veloce

Ordine di esecuzione delle operazioni.

In un sistema a tre indirizzi, l'esecuzione di un'operazione aritmetica si può scomporre nei passi seguenti.

1. Trasferire il primo numero dalla memoria all'unità aritmetica; la sua cella è indicata dal primo indirizzo.

2. Trasferire il secondo numero, la cui cella è indicata dal secondo indirizzo.

3. Eseguire nell'unità aritmetica l'operazione specificata dal codice operativo.

4. Trasferire il risultato nella cella di memoria indicata dal terzo indirizzo.

5. Prelevare l'istruzione successiva e iniziarne l'esecuzione.

Il codice dell'istruzione viene accolto nel registro dell'istruzione, indicato nello schema con IMB (figura 4). Un commutatore elettronico, EC, decodifica il codice operativo e attiva la linea corrispondente all'operazione. Tramite l'unità di controllo CU, questo segnale predispone i circuiti necessari. Le sigle inglesi sono conservate per permettere il confronto con lo schema originale.

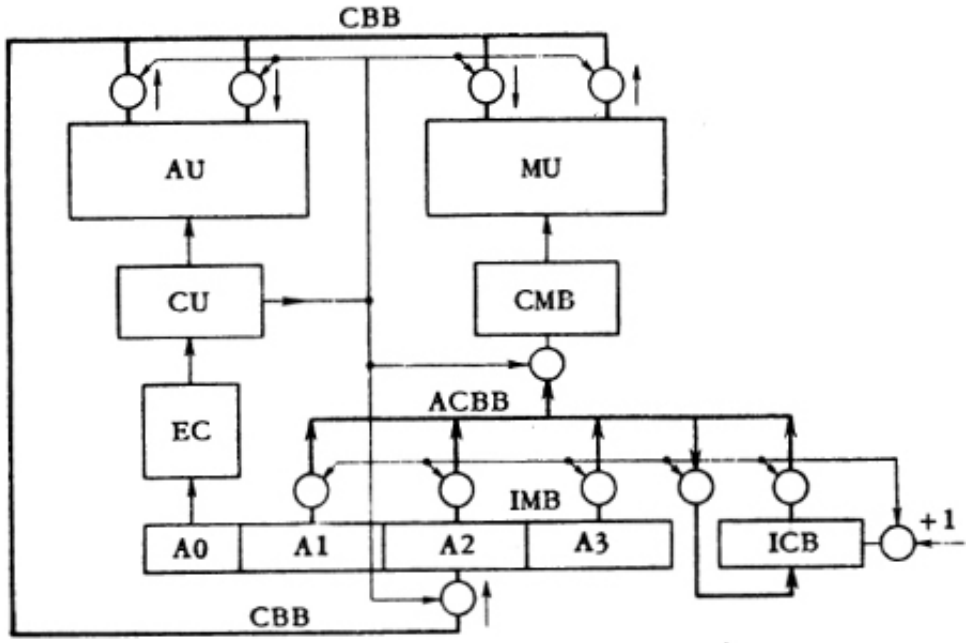


Figura 4. Schema strutturale di un calcolatore. AU: unità aritmetica; MU: memoria; CU: controllo; EC: commutatore elettronico; CMB: controllo della memoria; IMB: registro dell'istruzione; ICB: contatore delle istruzioni; CBB: linee dei dati; ACBB: linee degli indirizzi; A0: codice operativo; A1-A3: indirizzi.

Per prelevare il primo numero, il primo indirizzo A1 passa dall'IMB al blocco di controllo della memoria CMB attraverso le linee degli indirizzi ACBB. Il trasferimento è comandato dalla CU. Dalla cella corrispondente della memoria MU si preleva il numero, che passa attraverso le linee dei dati CBB all'unità aritmetica AU. Un segnale della CU abilita i circuiti d'ingresso dell'unità aritmetica.

Il secondo numero viene prelevato allo stesso modo: la CU trasferisce A2 dall'IMB al CMB; il numero selezionato nella memoria MU raggiunge l'AU attraverso le linee CBB.

L'unità aritmetica esegue sui due numeri l'operazione già specificata dal codice operativo.

Per registrare il risultato, il terzo indirizzo A3 passa dall'IMB al CMB attraverso ACBB, su comando della CU. Viene selezionata la cella corrispondente e abilitato il suo ingresso. La CU stabilisce se effettuare lettura o scrittura e trasferisce il risultato dall'AU alle linee CBB, dalle quali il valore viene registrato nella cella scelta.

La scelta delle istruzioni è affidata al blocco ICB, che contiene il numero dell'istruzione corrente. Normalmente basta aumentarlo di uno, mediante il circuito +1 comandato dalla CU. Il nuovo indirizzo passa dall'ICB al CMB attraverso ACBB; la memoria MU fornisce l'istruzione, che viene trasferita attraverso CBB nell'IMB. Si conclude così un ciclo, e il successivo esegue la nuova istruzione. L'ordine normale può essere modificato da un'istruzione di controllo: nel confronto, se il primo numero è minore del secondo si salta all'istruzione indicata dal terzo indirizzo; altrimenti si prosegue con quella successiva.

Quando l'IMB riceve un'istruzione di confronto, il commutatore EC ne decodifica il codice e attiva la linea che predispone i circuiti per tale operazione.

I due numeri da confrontare vengono prelevati come per un'operazione aritmetica. L'AU può confrontarli sottraendo il secondo dal primo. Se il risultato è negativo, la CU trasferisce A3 all'ICB attraverso ACBB; se è nullo o positivo, incrementa di uno l'ICB. Stabilito l'indirizzo successivo, l'istruzione viene prelevata dalla memoria nel modo consueto.

Unità aritmetica e unità di controllo.

I circuiti elettronici operano essenzialmente secondo il principio «sì o no»: il segnale è presente oppure assente. Entro margini appropriati, variazioni dei parametri fisici dei circuiti non modificano quindi l'informazione rappresentata.

Uno dei componenti più usati è il circuito bistabile, detto anche flip-flop. Il modello più semplice della figura 5 comprende due amplificatori a valvole, con resistori anodici R_a, collegati tramite i partitori R_1 e R_2. La polarizzazione O_B è scelta in modo che una valvola conduca e l'altra sia interdetta. Per la simmetria, sono possibili due stati stabili. Se la valvola sinistra è interdetta e la destra conduce, l'anodo sinistro O_1 ha tensione alta e quello destro O_0 tensione bassa, a causa della caduta sul resistore anodico. I partitori trasferiscono queste tensioni alle griglie opposte: la griglia sinistra resta a tensione bassa e la destra alta. Con parametri opportuni, questa retroazione mantiene lo stato.

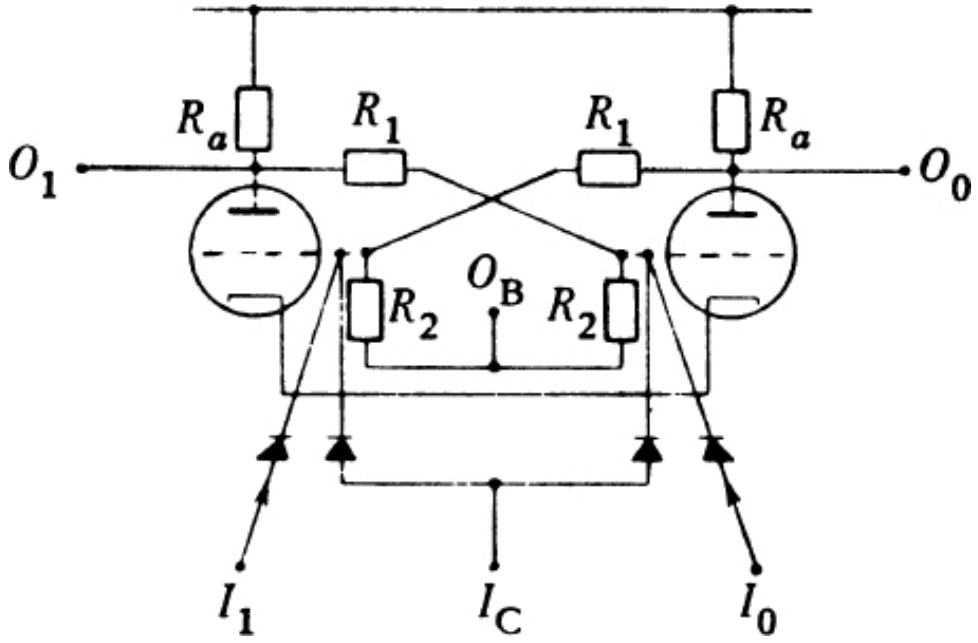


Figura 5. Circuito di un bistabile a valvole.

Viceversa, se la valvola sinistra conduce e la destra è interdetta, sono basse le tensioni dell'anodo sinistro e della griglia destra, alte quelle dell'anodo destro e della griglia sinistra.

Il passaggio fra i due stati si ottiene applicando impulsi negativi alle griglie attraverso diodi. Un impulso sulla griglia sinistra interdice la valvola: la sua tensione anodica sale, innalza la tensione della griglia destra e porta in conduzione la valvola destra. Si stabilisce così lo stato con tensione alta su O_1. Un impulso sulla griglia destra produce lo stato opposto. Nel circuito di conteggio opportunamente predisposto, un impulso applicato contemporaneamente ai due ingressi provoca il passaggio da uno stato all'altro.

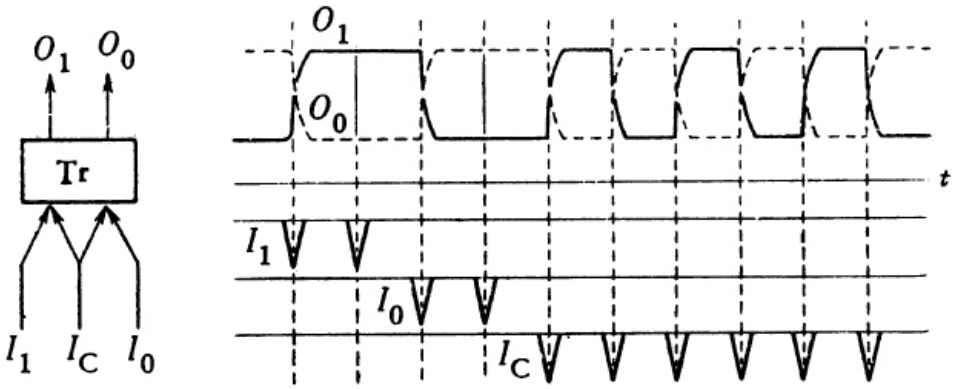


Figura 6. Funzionamento di un bistabile: ingressi, uscite e andamento temporale dei segnali.

Considerando ingressi le linee degli impulsi e uscite le tensioni anodiche, il funzionamento del bistabile si rappresenta come nella figura 6.

A uno stato si assegna il valore binario 0, per esempio quello con tensione alta sull'uscita destra O_0; all'altro il valore 1, con tensione alta sull'uscita sinistra O_1. Gli ingressi sono indicati con I_0, I_1 e I_C, ingresso di conteggio.

I bistabili possono conservare temporaneamente codici, formando registri di ricezione (figura 7). Un impulso negativo I_E applicato a tutti gli ingressi zero azzera il registro. Il codice di un numero o di un'istruzione viene poi applicato agli ingressi uno: le posizioni che ricevono l'impulso passano a 1 e vi rimangono fino al successivo azzeramento. Registri di questo tipo conservano operandi, istruzioni, indirizzi di memoria e altre informazioni.

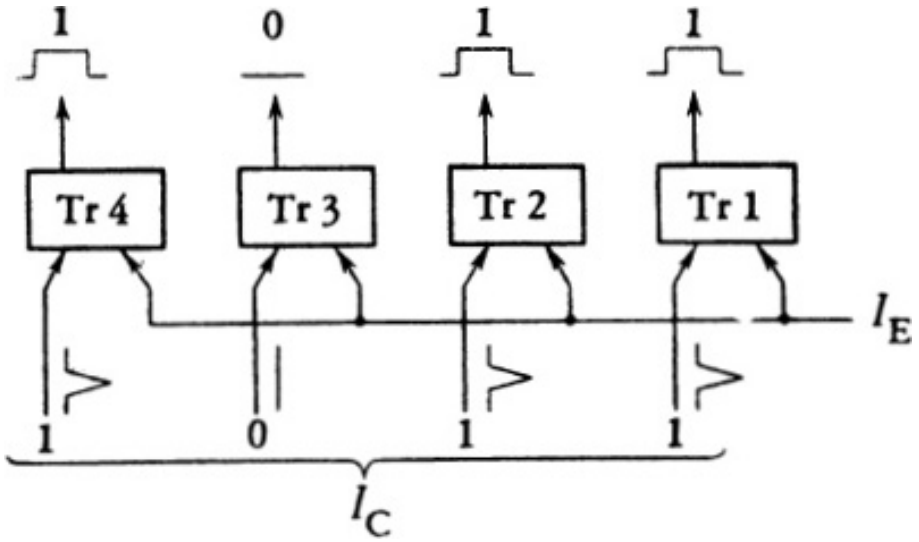


Figura 7. Registro di ricezione a bistabili. Tr indica un bistabile; I_E è il comando di azzeramento.

Un'altra applicazione riguarda i circuiti sommatori. Un impulso all'ingresso di conteggio cambia lo stato: da 0 a 1, oppure da 1 a 0; in assenza d'impulso lo stato resta invariato. Lo stato iniziale rappresenta una cifra del primo numero, mentre presenza o assenza d'impulso rappresenta la cifra corrispondente del secondo. Si realizzano così le regole 0 + 0 = 0, 0 + 1 = 1, 1 + 0 = 1 e 1 + 1 = 10. Nell'ultimo caso si scrive zero nella posizione corrente e si deve riportare uno alla successiva. Il riporto nasce dal passaggio del bistabile da 1 a 0: la tensione

dell'uscita sinistra scende e, attraversando un circuito con condensatore e resistore, genera un impulso negativo. Una linea di ritardo lo invia all'ingresso di conteggio della posizione successiva.

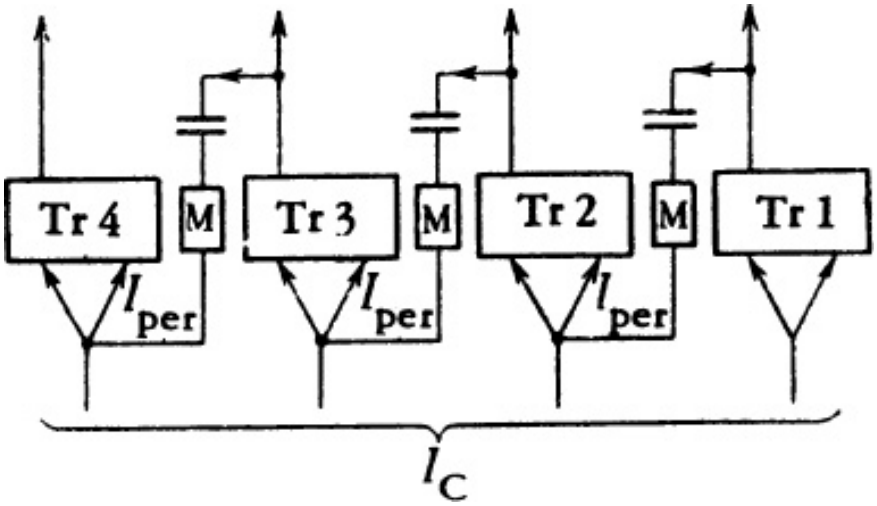


Figura 8. Sommatore a bistabili. M indica il ritardo; I_{per} è l'impulso di riporto.

La figura 8 mostra un semplice sommatore a bistabili. Prima si azzerano tutte le posizioni; poi gli impulsi del primo numero impostano il registro. All'arrivo del secondo numero avviene la somma cifra per cifra. Le posizioni che sommano due unità producono riporti, applicati dopo un ritardo t_d alle posizioni superiori. Questi possono a loro volta provocare nuovi riporti. Nel caso sfavorevole, una catena di cifre uguali a uno propaga il riporto lungo tutte le posizioni, richiedendo un tempo dell'ordine del ritardo elementare moltiplicato per il numero di cifre. Circuiti più complessi evitano questa propagazione interamente successiva e accelerano la somma.

Per moltiplicare, l'unità della figura 9 comprende due registri, R_1 e R_2, per moltiplicando e moltiplicatore, e un sommatore, Add. Si fa scorrere il moltiplicatore di una posizione verso destra. Se la cifra espulsa è 1, il relativo impulso abilita il trasferimento del moltiplicando al sommatore, attraverso il comando +N. Il prodotto parziale viene poi spostato di una posizione verso destra e si ripete il procedimento. Si accumula così la somma dei prodotti parziali. Servono tante ripetizioni quante sono le cifre. Il prodotto di due numeri di n cifre può occupare $2n$ cifre: le n più alte restano nel sommatore, mentre le più basse entrano progressivamente nelle posizioni liberate del registro del moltiplicatore. Alla fine quest'ultimo contiene le n cifre inferiori del prodotto. In questo schema la moltiplicazione richiede approssimativamente il tempo di un'addizione moltiplicato per n.

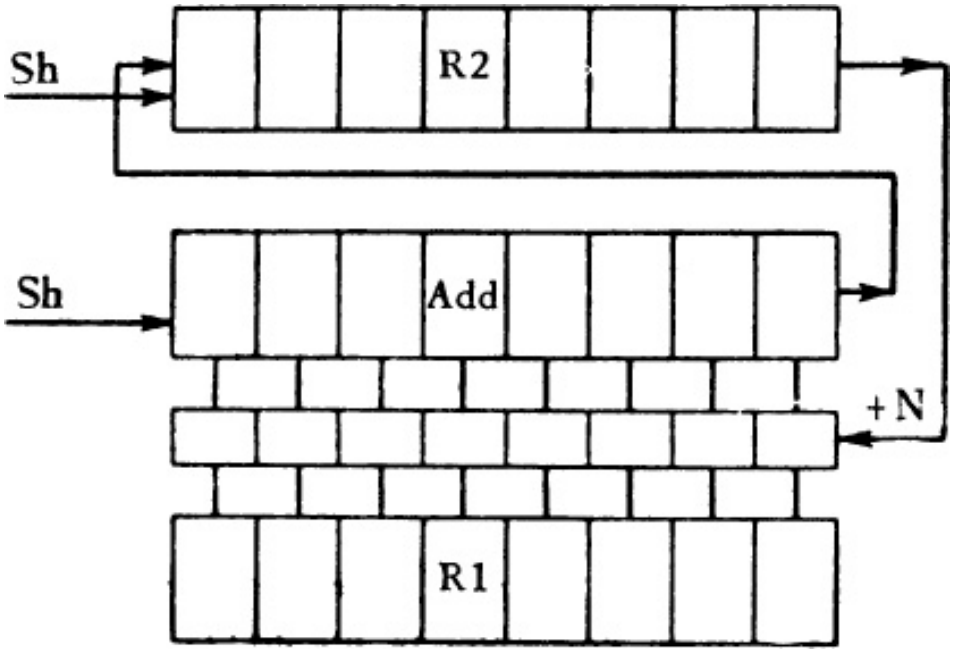


Figura 9. Moltiplicatore a bistabili. R1 e R2 sono i registri; Add è il sommatore; Sh indica lo scorrimento; +N comanda l'aggiunta del moltiplicando.

Lo scorrimento del codice si realizza con il circuito della figura 10. L'impulso I_{sh} azzera i bistabili; quelli precedentemente a 1 producono impulsi che, dopo il ritardo t_d, impostano a 1 i bistabili adiacenti. Ogni impulso di scorrimento sposta così il codice di una posizione.

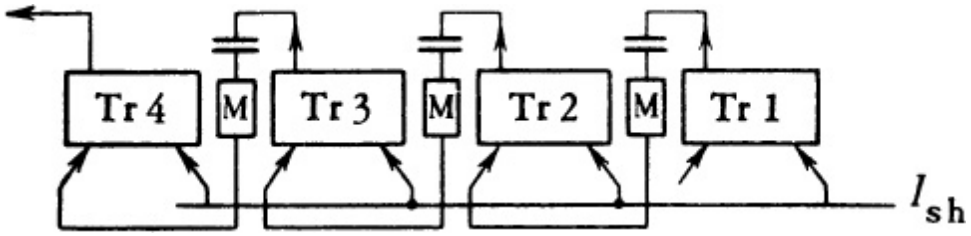


Figura 10. Circuito per lo scorrimento di un codice. I_{sh} è l'impulso di scorrimento e M il ritardo.

Un'unità con due registri e un sommatore permette anche di dividere un numero per un altro.

Di solito la si costruisce in modo da eseguire l'insieme delle operazioni aritmetiche e logiche richieste.

I bistabili servono anche a contare impulsi, funzione necessaria in diversi dispositivi di controllo. Il contatore della figura 11 differisce dal semplice sommatore della figura 8 essenzialmente per l'assenza delle linee di ritardo nei collegamenti dei riporti. Con n posizioni rappresenta gli stati da 0 a $2^n - 1$ e, dopo 2^n impulsi, ritorna allo stato iniziale. Modificando il circuito si può ottenere un contatore con un periodo diverso da una potenza di due.

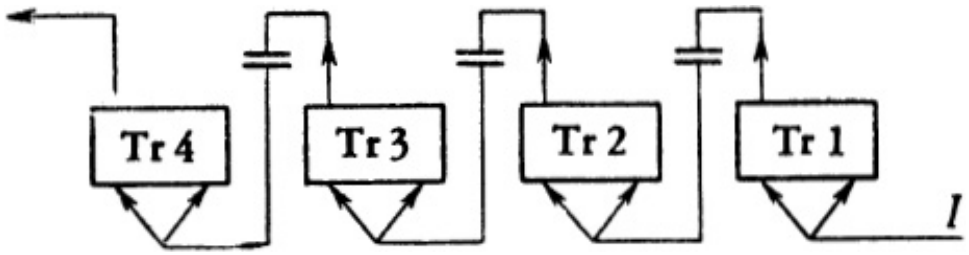


Figura 11. Contatore elettronico a bistabili.

Per realizzare operazioni logiche e circuiti di controllo si usano elementi di coincidenza, detti AND, invertitori e circuiti a diodi per la disgiunzione, detti OR.

Una porta AND produce un segnale soltanto quando tutti gli ingressi sono attivi: realizza il principio «l'uno e l'altro». Un invertitore realizza la negazione: ingresso attivo significa uscita inattiva, e viceversa. Una porta OR produce un segnale quando è attivo almeno uno degli ingressi, anche se ne sono attivi più di uno: si tratta della disgiunzione inclusiva.

Le porte AND sono usate per instradare i segnali nei circuiti desiderati. La figura 12 mostra la linea dei dati relativa a una sola cifra. Attraverso porte AND, essa è collegata agli ingressi e alle uscite delle celle di memoria, agli ingressi dei due registri dell'unità aritmetica e all'uscita del sommatore. Abilitando la porta di uscita di una cella, si pone sulla linea il bit conservato in quella cella; abilitando contemporaneamente l'ingresso di un registro, vi si trasferisce il bit. Analogamente, abilitando l'uscita del sommatore e l'ingresso di una cella, si registra in memoria il risultato. Prima di caricare un nuovo codice bisogna, nei registri qui descritti, cancellare quello precedente.

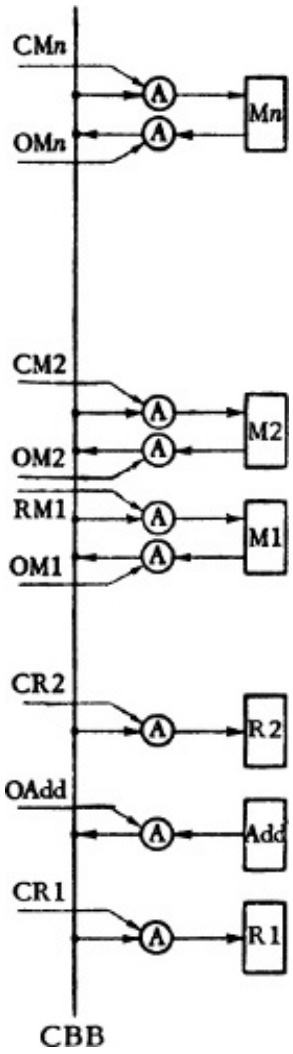


Figura 12. Instradamento dei segnali mediante porte AND. La linea comune collega celle di memoria, registri e sommatore; i segnali di comando abilitano i collegamenti necessari.

Questi esempi non esauriscono gli impieghi delle porte AND per instradare i segnali. Esse trovano applicazione nella memoria, nell'unità aritmetica e in quella di controllo.

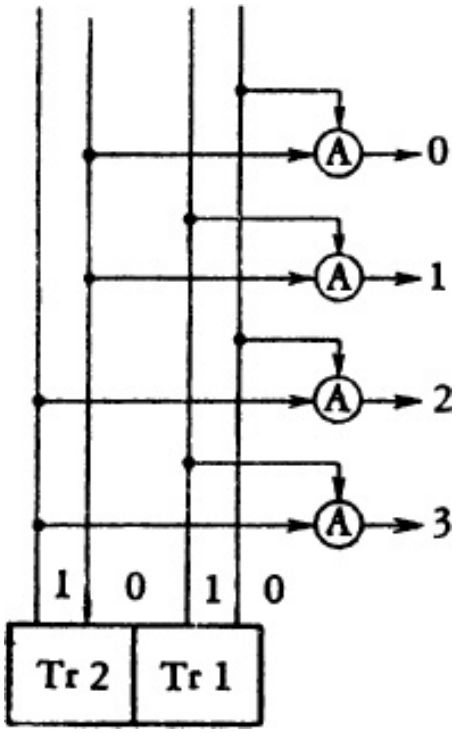


Figura 13. Commutatore elettronico con quattro linee d'uscita. A indica una porta AND; Tr1 e Tr2 sono i bistabili dell'indirizzo.

Le porte AND svolgono anche funzioni più complesse. Quando si accede alla memoria occorre spesso trasformare un indirizzo binario nell'attivazione della corrispondente cella. A questo provvede un commutatore elettronico, cioè un decodificatore. La figura 13 mostra il caso di quattro linee d'uscita, selezionate mediante due bistabili. Le quattro combinazioni possibili sono riportate nella tavola 5.

Tavola 5. Stati dei due bistabili e tensioni delle uscite

Codice	2° bistabile uscita sinistra	2° bistabile uscita destra	1° bistabile uscita sinistra	1° bistabile uscita destra
00	L	H	L	H
01	L	H	H	L
10	H	L	L	H
11	H	L	H	L

L: tensione bassa; H: tensione alta.

Se il livello attivo è quello alto, per selezionare la linea zero si collegano gli ingressi della porta AND alle uscite destre di entrambi i bistabili: la porta sarà attiva soltanto per il codice 00. Per la linea uno, codice 01, si collegano l'uscita sinistra del primo bistabile e quella destra del secondo. Lo stesso criterio determina i collegamenti delle linee due, codice 10, e tre, codice 11.

Porte AND, invertitori e porte OR possono formare anche circuiti aritmetici. Per sommare due cifre binarie si hanno le quattro combinazioni della tavola 6.

Tavola 6. Somma di due cifre binarie

N.	1° addendo	2° addendo	Somma	Riporto alla posizione superiore
1	0	0	0	0
2	0	1	1	0
3	1	0	1	0
4	1	1	0	1

Il circuito della figura 14 realizza queste relazioni ed è detto semisommatore. La porta AND genera il riporto quando entrambi gli ingressi valgono uno, cioè nella combinazione 4. Per la cifra della somma, combinazioni 2 e 3, deve essere attivo almeno un ingresso ma non il riporto: la condizione si realizza con AND, invertitore e OR a diodi. Per sommare numeri occorre considerare anche il riporto proveniente dalla posizione inferiore. Lo si aggiunge alla somma delle due cifre mediante un secondo semisommatore; i due segnali di riporto vengono riuniti. Si ottiene così un sommatore completo per una posizione.

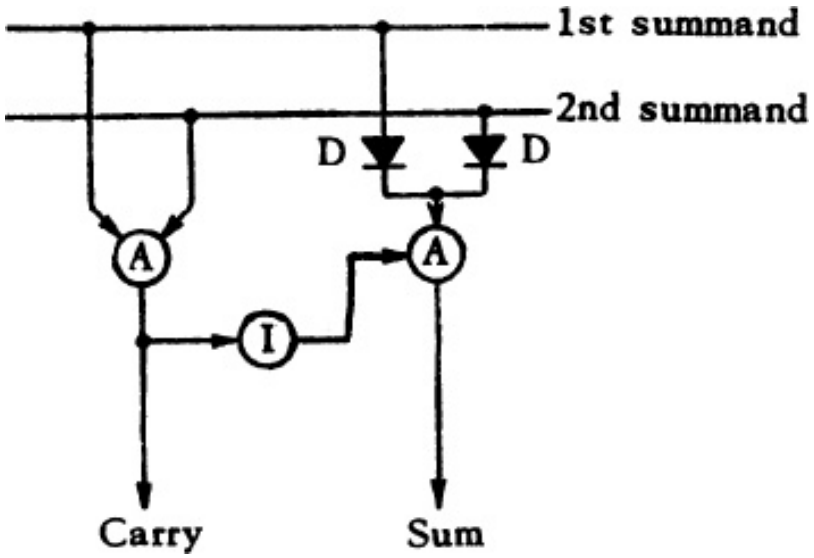


Figura 14. Semisommatore a una cifra. 1st summand e 2nd summand: primo e secondo addendo; Carry: riporto; Sum: somma; A: AND; I: invertitore; D: diodo.

Il circuito del sommatore completo può anche essere costruito direttamente, analizzando tutte le combinazioni dei due bit e del riporto entrante.

Questi circuiti sono particolarmente adatti alle macchine seriali. Le cifre viaggiano su una sola linea a intervalli rigorosamente determinati. Per sommare due numeri basta allora un sommatore a una cifra (figura 15): i due codici arrivano ai suoi ingressi cominciando dalle cifre meno significative. L'uscita del riporto è collegata al terzo ingresso mediante una linea di ritardo pari all'intervallo fra due cifre. Un riporto viene così applicato esattamente quando

arrivano i bit della posizione successiva. La somma richiede il tempo necessario a far passare un intero codice numerico.

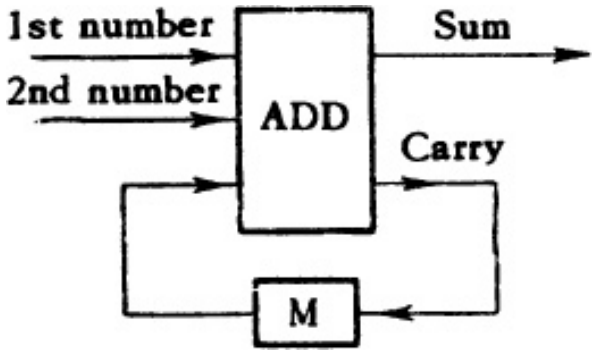


Figura 15. Sommatore seriale. 1st number e 2nd number: primo e secondo numero; Sum: somma; Carry: riporto; ADD: sommatore; M: linea di ritardo.

Anche la moltiplicazione seriale può usare un sommatore a una cifra, facendo passare i numeri tante volte quante sono le posizioni del codice. In questo schema richiede quindi circa *n* volte il tempo dell'addizione.

Unità di memoria.

La capacità della memoria, cioè quanti numeri si possono conservare, determina in larga misura le possibilità del calcolatore. Per le macchine universali dell'epoca vengono qui indicati valori usuali da 500 a 4.000 numeri.

I codici si potrebbero conservare in bistabili; nelle tecnologie descritte, però, l'apparecchiatura necessaria per una memoria completa sarebbe così grande da rendere questa soluzione poco usata.

Nelle macchine seriali ebbero larga diffusione le memorie elettroacustiche a tubi di mercurio (figura 16). Un impulso elettrico applicato a un cristallo di quarzo all'ingresso viene trasformato in una vibrazione ultrasonora, che si propaga lungo il tubo. Un secondo quarzo, all'estremità opposta, riconverte la vibrazione in impulso elettrico. Il segnale viene amplificato, ricondizionato e rimandato all'ingresso: i codici circolano così continuamente. Per scrivere nuovi numeri si introduce il codice dalla macchina e si interrompe per il tempo corrispondente il circuito di ricircolo. Per leggere un numero si abilita l'uscita quando il suo codice raggiunge la fine del tubo. Appositi circuiti eseguono automaticamente queste operazioni. Per ridurre l'apparecchiatura, ogni tubo contiene più numeri; l'accesso richiede perciò un'attesa, tanto maggiore quanto più lunga è la sequenza memorizzata.

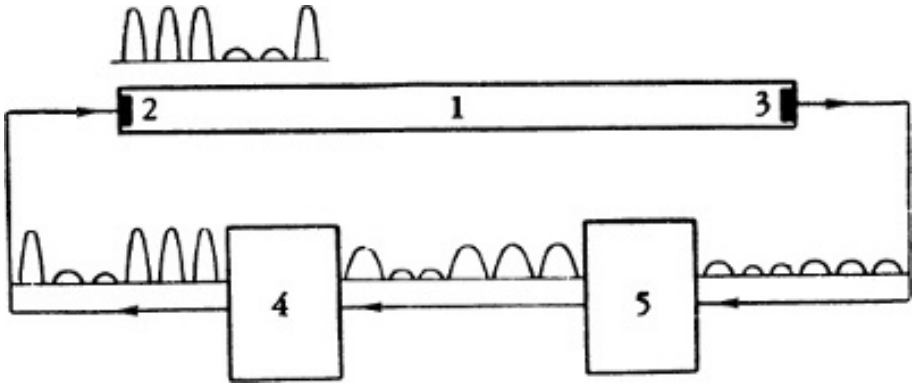


Figura 16. Conservazione dinamica di un codice in un tubo elettroacustico: 1, tubo al mercurio; 2, quarzo trasmittente; 3, quarzo ricevente; 4 e 5, stadi del circuito di ricircolo. Le forme d'onda mostrano gli impulsi trasmessi e ricevuti.

Le macchine seriali con queste memorie operavano tipicamente a 1.000-2.000 operazioni al secondo.

Un altro principio di memoria è la registrazione magnetica dei segnali, analoga a quella del suono. Si può usare un nastro oppure un tamburo in rotazione continua rivestito di materiale ferromagnetico (figura 17). Lungo una generatrice del tamburo sono disposte testine magnetiche. Impulsi di corrente negli avvolgimenti registrano segnali come magnetizzazione residua in punti determinati della superficie. Durante la rotazione, queste zone passano sotto le testine e vi inducono segnali elettrici, che vengono amplificati e trasmessi alle altre unità.

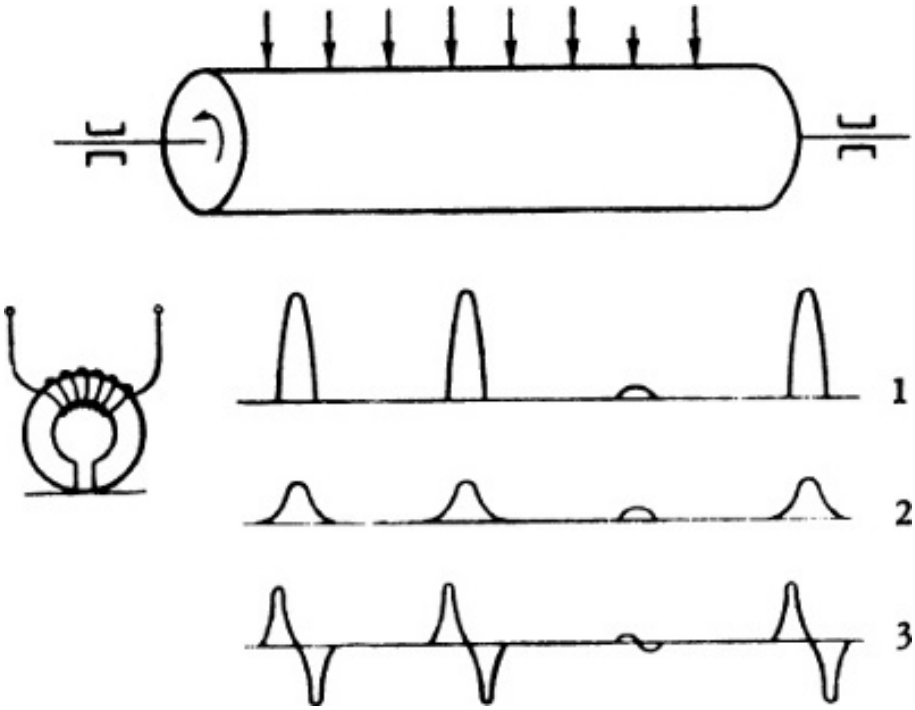


Figura 17. Principio della memoria a tamburo magnetico: 1, corrente nell'avvolgimento; 2, magnetizzazione residua; 3, forza elettromotrice indotta durante la lettura.

Il tamburo può servire sia sistemi seriali sia paralleli. L'attesa per raggiungere il dato, già presente nei tubi al mercurio, è però ancora più rilevante: per questo le memorie a tamburo

erano impiegate in macchine relativamente lente, dell'ordine di alcune centinaia di operazioni al secondo. D'altra parte aumentavano molto la capacità con un incremento accettabile dell'apparecchiatura. Tamburi e nastri venivano quindi usati anche come memorie ausiliarie esterne, accanto alla memoria operativa veloce.

Nei calcolatori veloci paralleli si usavano anche memorie a tubi catodici (figura 18). Un fascio elettronico diretto verso un punto dello schermo vi produce una carica che persiste abbastanza a lungo da rappresentare un bit. Per leggere si dirige nuovamente il fascio sul punto: se l'elemento non era carico, si carica e produce un impulso attraverso la piastra di segnale e l'amplificatore; se era già carico, la risposta è diversa. Si distingue così lo stato memorizzato. Dopo la lettura bisogna ripristinare lo stato precedente mediante un circuito automatico. Occorre inoltre rinnovare periodicamente le registrazioni per compensare elettroni dispersi e perdite attraverso il dielettrico.

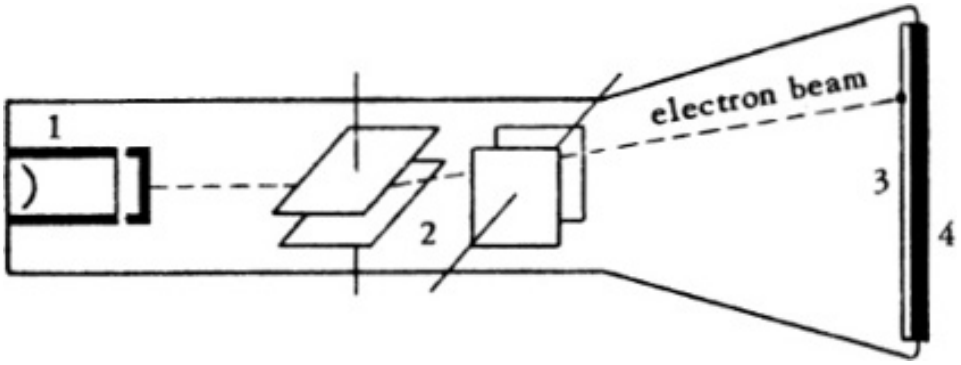


Figura 18. Memoria a tubo catodico: 1, sorgente di elettroni; 2, piastre di deflessione; 3, schermo; 4, piastra di segnale. Electron beam: fascio elettronico.

Sullo schermo si disponevano normalmente 1.024 punti, in una matrice 32 × 32, oppure 2.048 punti, in una matrice 32 × 64. Due coppie di piastre di deflessione, comandate da opportune tensioni, dirigevano il fascio sul punto richiesto.

Nei sistemi paralleli ogni posizione binaria richiedeva un tubo distinto; l'accesso alle cifre di un numero avveniva simultaneamente su tutti i tubi. Il ciclo di accesso poteva ridursi a pochi microsecondi.

All'epoca del testo si stavano diffondendo memorie a elementi magnetici con ciclo d'isteresi rettangolare (figura 19). Un impulso di corrente in un verso magnetizza il nucleo positivamente, un impulso opposto lo magnetizza negativamente.

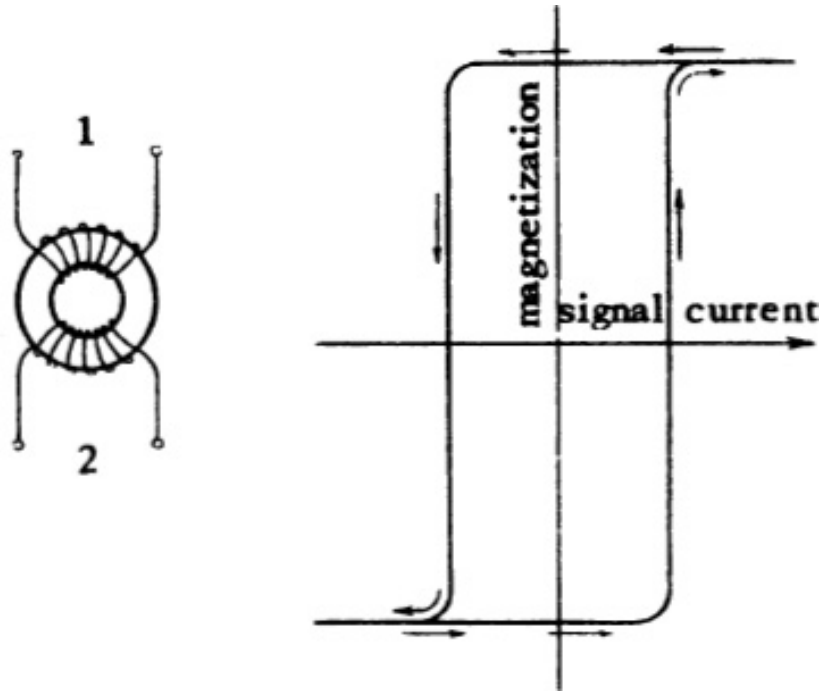


Figura 19. Elemento di memoria con ciclo d'isteresi rettangolare: 1, avvolgimenti d'ingresso; 2, avvolgimento d'uscita. Magnetization: magnetizzazione; signal current: corrente del segnale.

Cessato l'impulso, il nucleo conserva la magnetizzazione, che rappresenta il bit. Per leggere si applica un impulso di polarità fissata, per esempio positiva. Se il nucleo era magnetizzato negativamente, cambia stato e la variazione del flusso induce una forza elettromotrice nell'avvolgimento di lettura, collegato a un amplificatore. Se era già positivo, non si produce il medesimo cambiamento. Si riconosce così il valore registrato. Dopo questa lettura occorre ripristinare lo stato originario, mediante un circuito apposito.

§4. Prospettive di sviluppo e impiego dei calcolatori elettronici

L'uso dei calcolatori è destinato, secondo la prospettiva degli autori, a influire profondamente su molti settori scientifici e tecnici, soprattutto sulle scienze fisiche e matematiche. È quindi opportuno delineare le prospettive che si aprivano allora e il loro significato per la matematica.

Ampliamento delle applicazioni: 1) Macchine migliori.

La produzione di calcolatori veloci era in pieno sviluppo: si perfezionavano le strutture, si introducevano nuovi principi fisici e nuove combinazioni di componenti. Ci si attendevano maggiore velocità, capacità di memoria, regolarità e affidabilità, insieme a una semplificazione della costruzione e dell'uso che ne favorisse la diffusione.

Un altro fattore era la varietà dei modelli. Accanto alle grandi macchine sarebbero comparsi calcolatori piccoli, semplici da usare e accessibili a istituti scientifici, uffici di progettazione e

fabbriche. Accanto ai modelli universali, macchine specializzate più semplici avrebbero affrontato classi determinate di problemi. Si sviluppavano inoltre sistemi che ricevevano dati da dispositivi fisici, li elaboravano numericamente e restituivano risultati in forma continua, come curve o parametri di comando.

2. Migliore programmazione.

Per problemi semplici è agevole costruire programmi nel modo descritto nel §2; nei problemi reali di grandi dimensioni il lavoro diventa complesso e minuzioso. Una biblioteca di sottoprogrammi per funzioni fondamentali e operazioni come inversione di matrici o integrazione numerica lo alleggerisce. Restano però difficili il collegamento dei sottoprogrammi, l'assegnazione e modifica degli indirizzi, le prove e la riorganizzazione del programma. Questo lavoro, che richiede esperienza, può ritardare molto l'introduzione di nuovi problemi sulla macchina.

Gli autori indicano due direzioni di sviluppo. La prima consiste nel far costruire automaticamente il programma alla macchina stessa: formule e struttura logica del problema vengono introdotte in forma codificata e un apposito «programma di programmazione» le trasforma nel programma da eseguire.

La seconda consiste nell'impiegare un programma universale che interpreti direttamente un piano generale di calcolo. Questo piano può indicare operazioni complesse, come la soluzione di un sistema di equazioni; il programma universale ne cura l'esecuzione e l'associazione dei risultati ai singoli compiti, senza richiedere ogni volta la scrittura di tutte le istruzioni elementari.

3. Compiti intellettuali più complessi.

Un'ulteriore prospettiva è l'uso delle macchine non soltanto per calcoli numerici, ma anche per operazioni standard dell'analisi.

In alcuni casi ciò è già realizzabile in modo diretto. Rappresentando un polinomio mediante i suoi coefficienti, moltiplicazione e divisione fra polinomi diventano operazioni aritmetiche su successioni, facilmente programmabili. Con una codifica appropriata delle funzioni, un programma può produrre la derivata simbolica di una funzione elementare nella medesima codifica. Si apre così la possibilità di risolvere problemi con un metodo assegnato - per esempio un sistema differenziale mediante serie di potenze - eseguendo automaticamente sia i passaggi analitici sia quelli numerici. Le macchine possono dunque svolgere compiti intellettuali sottili, purché formalizzati in procedure, così come le macchine industriali sostituiscono non soltanto il lavoro dello scaricatore, ma anche quello della sarta.

Influenza sui metodi numerici e approssimati.

Gli strumenti impiegati influenzano i metodi di lavoro. Per esempio, formule trigonometriche predisposte per il calcolo con logaritmi non sono convenienti su una macchina che esegue direttamente moltiplicazioni e divisioni. Già l'uso di una calcolatrice da tavolo richiedeva schemi diversi nei metodi approssimati, per esempio procedimenti per equazioni differenziali che non facessero uso di tavole di differenze.

Il mutamento radicale degli strumenti di calcolo ha quindi modificato il modo di valutare non soltanto i metodi numerici, ma anche molti problemi matematici e le loro applicazioni.

Consideriamo alcune questioni nelle quali questo cambiamento è particolarmente evidente.

Tavole e altre rappresentazioni delle funzioni.

I calcolatori hanno trasformato la produzione di tavole: invece di singole opere, sono diventate possibili centinaia di tavole all'anno, con valori accurati delle principali funzioni speciali, anche di più variabili. Ma cambia anche la struttura più conveniente. Per una macchina veloce possono essere preferibili tavole compatte, con valori di base distanziati e interpolazione di ordine elevato.

Spesso conviene sostituire le tavole con altre rappresentazioni: polinomi di migliore approssimazione su sottointervalli, sviluppi in frazioni continue, formule basate sul calcolo numerico di un integrale che rappresenta la funzione e così via. Ciascuna può essere inserita vantaggiosamente, nei casi appropriati, nel programma di calcolo della funzione.

Funzioni speciali e soluzioni analitiche particolari.

Le funzioni speciali e le espressioni analitiche finite conservano importanza per lo studio qualitativo e per riconoscere le singolarità, aspetti rilevanti anche per la soluzione numerica. In alcuni problemi di grandi dimensioni possono offrire il procedimento numerico più economico. In molti casi particolari, però, costruire espressioni complesse mediante funzioni speciali introdotte per facilitare il calcolo non risulta più conveniente: i metodi numerici generali possono essere più semplici e rapidi.

Gli autori osservano perciò che molti sforzi dedicati a esprimere in forme analitiche complicate soluzioni di particolari problemi tecnici e meccanici possono perdere la loro originaria utilità computazionale.

Scelta dei metodi di calcolo.

La grande potenza delle macchine non rende inutile migliorare i metodi approssimati, né autorizza a usare sempre quelli più rudimentali. Soltanto nei problemi unidimensionali più semplici, che richiedono comunque poche migliaia di passi, il calcolo può terminare in pochi secondi o minuti sulle macchine considerate.

Nei problemi più complessi i passi possono essere centinaia di milioni, e ridurne il numero è essenziale. È quindi di grande importanza pratica sviluppare metodi efficaci per problemi multidimensionali: interpolazione di funzioni di più variabili, integrali multipli, sistemi

algebrici o trascendenti non lineari, equazioni integrali tridimensionali, sistemi di equazioni alle derivate parziali e altri ancora.

Cambiano anche i criteri di valutazione. Contano la facilità di esecuzione sulla macchina e la generalità, cioè l'applicabilità a grandi famiglie di problemi. Perdono valore operativo i metodi che richiedono di sfruttare ogni volta peculiarità molto speciali o l'abilità personale di chi dirige il calcolo.

Assumono grande importanza i metodi generali: differenze finite, metodi variazionali, metodo del gradiente, iterazioni, linearizzazione e altri procedimenti analoghi.

Nella scelta bisogna però considerare anche la macchina concreta: utilizzare bene la memoria operativa, limitare i trasferimenti dall'esterno, prevedere controlli intermedi e facilitare la programmazione.

Non bisogna pensare che un calcolatore esegua soltanto schemi semplici e uniformi. La programmazione permette procedimenti complessi con numerose diramazioni, capaci di cambiare percorso secondo i risultati intermedi, anche quando ciò sarebbe difficile a mano. Occorre però descrivere preventivamente le regole che governano queste possibilità.

Né sono esclusi i metodi che richiedono operazioni algebriche o simboliche: come abbiamo visto, si possono programmare anche alcune operazioni dell'analisi.

Significato delle stime dell'errore.

Le stime asintotiche acquistano maggiore importanza perché le macchine veloci rendono praticabili grandi valori di n - per esempio molte equazioni nel sistema che approssima un'equazione integrale - e piccoli passi nei metodi alle differenze. Nel confronto fra metodi, le stime della rapidità di convergenza diventano quindi un criterio fondamentale.

Meritano particolare attenzione le stime a posteriori, calcolate a partire dalla soluzione approssimata già ottenuta. Inserite nel programma, possono guidare il seguito del calcolo: se l'errore è troppo grande, il programma può ripetere automaticamente il procedimento con passo dimezzato. Queste stime possono risultare più pratiche delle stime a priori, spesso più conservative o difficili da calcolare.

Possibilità di analisi teorica.

L'approssimazione numerica può fornire informazioni anche sull'esistenza e unicità della soluzione e su una regione che la contiene, quando viene accompagnata da opportune stime rigorose dell'analisi funzionale. Poiché una dimostrazione esclusivamente teorica può essere lunga, complessa o impraticabile, l'uso di risultati numerici verificati a questo scopo è di grande interesse.

Nuovi problemi dei metodi numerici.

L'aumento delle possibilità di calcolo ha introdotto nuove questioni. La soluzione di grandi sistemi lineari, un tempo occasionale, è divenuta una fase abituale di molti procedimenti. Diventa quindi essenziale sapere come gli arrotondamenti dei coefficienti e delle operazioni intermedie influenzino la precisione delle incognite. La questione ha dato origine a numerose ricerche.

L'integrazione numerica di sistemi differenziali su lunghi intervalli, con piccoli passi, ha reso cruciale la stabilità. L'esperienza di calcolo e le successive indagini teoriche hanno modificato profondamente il giudizio su vari metodi d'integrazione.

La stabilità è altrettanto importante per l'applicazione delle differenze finite alle equazioni alle derivate parziali.

Nuovi metodi.

Le macchine hanno favorito metodi nuovi e reso praticabili procedimenti antichi prima troppo laboriosi. Un esempio è il campionamento casuale, o metodo Monte Carlo. Si costruisce un problema probabilistico la cui probabilità o speranza matematica coincida con la quantità cercata, e la si stima mediante una media di prove. Per esempio, per trovare l'area di una figura definita da $F(x, y) \geqslant 0$ e contenuta nel quadrato unitario [0, 1] × [0, 1], si generano coppie (x, y) uniformemente distribuite nel quadrato, con prove indipendenti, e si calcola la frazione che soddisfa la disuguaglianza. A mano sarebbe un metodo poco efficiente; sulla macchina diventa praticabile. Si possono usare tavole di numeri casuali. Per certi problemi, come integrali di molte variabili senza esigenze di elevatissima precisione, può essere particolarmente efficace.

Metodi analoghi si applicano all'inversione di matrici mediante campioni organizzati in catene di Markov, oppure alle equazioni alle derivate parziali quando si dispone di un processo stocastico opportunamente collegato al problema.

Significato per analisi, meccanica e fisica.

Lo studio dei problemi multidimensionali che conducono a equazioni integrali e problemi al contorno della fisica matematica acquista maggiore interesse pratico. Le nuove tecniche rendono attuabili metodi prima troppo onerosi, proprio mentre la soluzione di tali problemi diventa sempre più urgente.

Anche questi metodi devono essere valutati in base alla possibilità concreta di applicarli.

La capacità di compiere molte prove numeriche con precisione sufficiente amplia enormemente l'uso degli «esperimenti matematici» nelle indagini preliminari. Diventa importante sviluppare sia le idee generali della sperimentazione numerica e del campionamento, sia applicazioni specifiche, per esempio allo studio qualitativo delle equazioni differenziali.

Il calcolo automatico può servire anche a questioni puramente teoriche: per esempio a migliorare le costanti di disuguaglianze e stime nell'analisi funzionale. Impieghi di questo genere si incontrano anche nella teoria dei numeri.

Si può inoltre verificare la correttezza di formule e deduzioni della logica matematica. Poiché molte proposizioni e dimostrazioni possono essere formalizzate, diventa possibile affidare alla macchina il controllo logico di deduzioni esplicitamente codificate.

In meccanica e fisica, il primo effetto è l'ampliamento dell'applicazione della matematica. In precedenza il volume e la complessità dei calcoli limitavano fortemente la trattazione dei problemi concreti: un solo problema poteva richiedere mesi o anni di lavoro. Benché formulazioni generali e metodi teorici fossero noti, soluzioni effettive erano spesso disponibili soltanto per casi idealizzati: problemi piani o assialsimmetrici, contorni molto semplici, ali di lunghezza infinita.

Le soluzioni matematiche servivano quindi soprattutto a indagini qualitative e preliminari, da completare con costosi esperimenti fisici, più che a fornire direttamente tutti i valori numerici necessari.

I nuovi strumenti aprono invece la possibilità di affrontare su vasta scala problemi spaziali, contorni complessi ed equazioni non lineari, avvicinandosi maggiormente alle condizioni reali.

Per sfruttare queste possibilità occorre sviluppare ancora i metodi numerici e la loro realizzazione. Gli autori richiamano però i successi già ottenuti nella soluzione di sistemi alle derivate parziali in meteorologia, dinamica dei gas, meccanica dei materiali granulari e altri settori.

La maggiore aderenza ai problemi fisici, unita alla rapidità e flessibilità del calcolo, rende possibile sostituire in molti casi alcuni esperimenti fisici con esperimenti numerici. Si prospetta così un rafforzamento dei metodi teorici e computazionali nello studio della fisica e della meccanica.

Significato per tecnica e industria.

La rapidità delle soluzioni numeriche permette un impiego più ampio dei metodi teorici nella scienza delle costruzioni, nell'elettrotecnica, nella radiotecnica, nell'utilizzazione dell'energia idraulica e in altri campi. Si possono ottenere risultati più accurati e utili, e affrontare problemi nei quali l'analisi matematica non era stata applicata.

Si apre anche un impiego diverso: confrontare possibili progetti o distribuzioni di risorse e impianti. Nell'organizzazione industriale esistono molte alternative per assegnare i compiti e stabilirne la successione; scegliere la più produttiva ed economica è difficile. Se si programma un esame sistematico delle varianti secondo criteri assegnati, la macchina può confrontarne centinaia di migliaia, ben oltre le possibilità dei metodi usuali.

In particolare, configurazioni di circuiti a relè e contatti possono essere analizzate e verificate con metodi di logica matematica eseguibili al calcolatore. Si possono così selezionare le varianti che soddisfano determinati criteri e scegliere la migliore fra esse.

Gli autori considerano molto promettente anche il controllo automatico industriale mediante calcolatori collegati a servomeccanismi e dispositivi di trasmissione. Introducendo i dati geometrici di un manufatto e un programma adeguato, la macchina può determinare i parametri che comandano una pressa e modificano il pezzo. La velocità consente di controllare più macchine contemporaneamente. Lo stesso principio interessa la guida automatica di veicoli, compresi razzi interplanetari: il programma può usare non soltanto i dati iniziali, ma anche gli aggiornamenti di posizione forniti dagli strumenti.

La costruzione e lo studio dei calcolatori, insieme alle loro applicazioni, aprono dunque un vasto campo di attività ai matematici. Il capitolo si conclude con la previsione che le macchine matematiche avrebbero avuto un ruolo decisivo nello sviluppo futuro della tecnica e della cultura.

Letture consigliate

F. L. Alt, *Electronic digital computers: their use in science and engineering*, Academic Press, New York, 1958.

A. D. Booth e K. H. V. Booth, *Automatic digital calculators*, Academic Press, New York, 1953.

P. von Handel (a cura di), *Electronic computers: fundamentals, systems and applications*, Prentice-Hall, Englewood Cliffs, N. J., 1961.

G. N. Lance, *Numerical methods for high-speed computers*, Iliffe and Sons, London, 1960.

D. D. McCracken, *Digital computer programming*, Wiley, New York, 1957.

J. J. Murray, *Mathematical machines. I, Digital computers; II, Analog devices*, Columbia University Press, New York, 1961.

J. von Neumann, *The computer and the brain*, Yale University Press, New Haven, Conn., 1958.

R. K. Richards, *Arithmetic operations in digital computers*, Van Nostrand, New York, 1955.

G. R. Stibitz e J. A. Larrivee, *Mathematics and computers*, McGraw-Hill, New York, 1957.

M. V. Wilkes, *Automatic digital computers*, Wiley, New York, 1956.

Note al capitolo XIV

[1] Nella pratica si usano normalmente formule più elaborate e accurate per risolvere le equazioni differenziali ordinarie.

[2] Nella colonna «Calcoli», i numeri fra parentesi indicano l'operazione il cui risultato va usato, mentre il pedice indica il passo. La prima riga moltiplica *ah* per il risultato dell'operazione 2 al passo precedente, $(2)_{k-1}$. La seconda somma a quel valore il risultato dell'operazione 1 al passo corrente, $(1)_k$.

All'inizio si usano i dati y_0 e x_0, rispettivamente al posto dei risultati precedenti delle operazioni 2 e 3.

[3] Una cella destinata ai valori intermedi del calcolo viene qui chiamata cella di lavoro.

Nota editoriale

La traduzione conserva le quattro sezioni, le diciannove figure, le sei tavole numerate, le formule, la bibliografia e le tre note. Le tavole dei programmi e dei circuiti sono ricomposte in italiano, insieme ai piccoli prospetti degli indirizzi e dei sistemi di numerazione. La figura 2, che rappresenta i campi di un'istruzione, è ricomposta tipograficamente; le altre figure restano quelle dell'edizione fornita, con didascalie e legende italiane. Le due parti della figura 3 sono riunite. Le sigle degli schemi originali sono conservate e spiegate.

Il capitolo descrive macchine e prospettive degli anni Cinquanta: velocità, capacità delle memorie, confronti fra architetture, valutazioni economiche e previsioni sono riferiti a quel contesto. I giudizi sulla creatività delle macchine e sull'utilità delle rappresentazioni analitiche sono presentati come posizioni degli autori. La precisione numerica non è letteralmente infinita; in virgola mobile il numero di cifre significative fornisce una precisione relativa dello stesso ordine, entro il campo rappresentabile, non un errore relativo identico per tutti i numeri.

Nel programma di Eulero il confronto finale è $x \geq x_n$; la tavola corregge il riferimento ambiguo dell'originale. Nella nota 2 i dati iniziali y_0 e x_0 corrispondono rispettivamente alle operazioni 2 e 3. Nel programma del polinomio, il trattino del secondo indirizzo nella prima istruzione indica una copia, equivalente a sommare zero. L'incremento del codice modifica soltanto il campo del secondo indirizzo; si presuppone una codifica compatibile e senza riporto verso altri campi. Si tratta del procedimento storico di codice automodificante descritto dagli autori.

Nell'esempio di allineamento degli esponenti, lo scarto delle ultime cifre richiede un segno di approssimazione: $10^2 \cdot 0{,}7587 + 0{,}3743 \approx 10^2 \cdot 0{,}7624$. La notazione binaria delle immagini conserva il punto separatore; nel codice a virgola mobile dell'esempio, il gruppo iniziale 0011 rappresenta l'esponente 3 e il gruppo successivo la parte significativa. Le verifiche mediante ripetizione, dimezzamento del passo o valori di riferimento forniscono controlli, ma non garantiscono incondizionatamente l'assenza di ogni errore. Analogamente, un risultato numerico può sostenere una prova di esistenza o unicità soltanto insieme a stime rigorose.

Per le valvole si usano i termini «in conduzione» e «interdetta», evitando l'ambiguità dell'inglese open/closed. Si correggono i refusi nella descrizione della somma di due bit e nella didascalia del contatore della figura 11. La porta OR è inclusiva; per il sommatore completo formato da due semisommatori si esplicita la riunione dei riporti. Un contatore a n bit ha stati da 0 a $2^n - 1$ e periodo 2^n. La didascalia della figura 16 distingue gli stadi di ricircolo dalle forme d'onda disegnate. Per l'esempio Monte Carlo si precisano distribuzione uniforme e indipendenza dei punti nel quadrato unitario.

CAPITOLO XV

Teoria delle funzioni di una variabile reale

S. B. Stečkin

§1. Introduzione

Tra la fine del XVIII e l'inizio del XIX secolo il calcolo differenziale e integrale era ormai sostanzialmente costituito. Fino ad allora, e in realtà per tutto il XIX secolo, i matematici si erano dedicati a costruirne i diversi rami, a scoprire nuovi risultati e ad ampliare i campi di applicazione ai problemi della meccanica, dell'astronomia e della tecnica. Era ormai possibile esaminare nel loro insieme i risultati ottenuti, studiarli sistematicamente e approfondire il significato dei concetti fondamentali dell'analisi. Divenne allora evidente che i fondamenti dell'analisi non erano del tutto soddisfacenti.

Già nel XVIII secolo i maggiori matematici non concordavano su che cosa fosse una funzione. Questa incertezza si manifestava in lunghe controversie sulla correttezza di una soluzione o di un risultato matematico concreto. A poco a poco divenne chiaro che occorreva precisare anche gli altri concetti fondamentali dell'analisi. Una comprensione insufficiente della continuità e delle proprietà delle funzioni continue aveva condotto ad affermazioni erronee, per esempio che una funzione continua fosse sempre derivabile. La matematica era giunta a trattare funzioni tanto complesse da non poter più fare affidamento sull'intuizione e sulle congetture. Nacque così la necessità effettiva di mettere ordine nei concetti fondamentali dell'analisi.

Il primo serio tentativo in questa direzione fu compiuto da Lagrange; Cauchy seguì poi la stessa strada. Cauchy precisò le definizioni di limite, continuità e integrale e ne diffuse l'uso nelle forme giunte fino a noi. Quasi contemporaneamente il matematico ceco Bolzano studiò rigorosamente le proprietà fondamentali delle funzioni continue.

Consideriamo più da vicino queste proprietà. Sia data una funzione continua $f(x)$ su un intervallo $[a, b]$, cioè per tutti i numeri che soddisfano le disuguaglianze $a \leqslant x \leqslant b$. In precedenza si considerava evidente che, se la funzione assume valori di segno opposto agli estremi dell'intervallo, essa deve annullarsi in qualche punto intermedio. Ora questo fatto ricevette una dimostrazione rigorosa. Analogamente si dimostrò che una funzione continua su un intervallo chiuso e limitato assume in certi punti il suo valore massimo e il suo valore minimo.

Lo studio di queste proprietà rese necessario approfondire la natura dei numeri reali. Nacque così la teoria dei numeri reali e vennero formulate con chiarezza le proprietà fondamentali della retta numerica.

Gli sviluppi successivi dell'analisi matematica richiesero lo studio di funzioni sempre più «irregolari», in particolare discontinue. Funzioni discontinue compaiono, per esempio, come limiti di funzioni continue, senza che si sappia a priori se la funzione limite sia continua, e nella schematizzazione di processi con variazioni brusche e improvvise. Si presentò dunque un nuovo compito: estendere gli strumenti dell'analisi alle funzioni discontinue.

Riemann studiò a quali classi di funzioni discontinue si potesse estendere il concetto di integrale. Da queste ricerche sui fondamenti dell'analisi nacque una nuova disciplina matematica: la teoria delle funzioni di una variabile reale.

Mentre l'analisi matematica classica tratta essenzialmente funzioni «regolari», per esempio continue o derivabili, la teoria delle funzioni di una variabile reale studia classi di funzioni molto più ampie. Se nell'analisi matematica un'operazione, per esempio l'integrazione, viene definita per le funzioni continue, la teoria delle funzioni di variabile reale cerca di stabilire a quali classi di funzioni tale definizione sia applicabile e come modificarla per ampliarne la portata. In particolare, soltanto questa teoria ha potuto rispondere in modo soddisfacente alla domanda su che cosa sia la lunghezza di una curva e per quali curve abbia senso parlare di lunghezza.

Il fondamento su cui si costruisce questa teoria è la teoria degli insiemi.

Cominceremo perciò con gli elementi della teoria degli insiemi; passeremo poi agli insiemi di punti e concluderemo illustrando uno dei concetti fondamentali della teoria delle funzioni di una variabile reale: l'integrale di Lebesgue.

§2. Insiemi

Gli uomini hanno continuamente a che fare con collezioni di oggetti. Come si è spiegato nel capitolo I, da ciò si è sviluppato il concetto di numero e, in seguito, quello di insieme, uno dei concetti matematici primitivi fondamentali, che non si presta a una definizione mediante concetti più semplici. Le osservazioni seguenti intendono illustrare che cosa sia un insieme, senza pretendere di definirlo.

Si chiama insieme un aggregato o una collezione di oggetti riuniti secondo un certo criterio o una certa regola. Il concetto di insieme nasce da un'astrazione. Nel considerare una collezione di oggetti come un insieme, prescindiamo dai legami e dalle relazioni fra i diversi oggetti che la compongono, conservandone però le caratteristiche individuali. Così l'insieme di cinque monete e quello di cinque mele sono insiemi diversi. Invece l'insieme di cinque monete disposte in cerchio e quello delle stesse monete disposte in fila sono un solo e medesimo insieme.

Facciamo alcuni esempi. Possiamo considerare i granelli di un mucchio di sabbia, l'insieme dei pianeti del sistema solare, quello delle persone presenti in una casa in un determinato istante o quello delle pagine di questo libro. In matematica incontriamo continuamente

insiemi: tutte le radici di una data equazione, tutti i numeri naturali, tutti i punti di una retta e così via.

La disciplina matematica che studia le proprietà generali degli insiemi, cioè quelle indipendenti dalla natura degli oggetti che li compongono, si chiama teoria degli insiemi. Il suo sviluppo rigoroso ebbe inizio tra la fine del XIX e l'inizio del XX secolo. Il fondatore della teoria scientifica degli insiemi è il matematico tedesco G. Cantor.

Le ricerche di Cantor nacquero dallo studio della convergenza delle serie trigonometriche. È un fenomeno frequente: lo studio di problemi matematici concreti conduce spesso alla costruzione di teorie molto astratte e generali. Il valore di queste costruzioni sta nel fatto che esse si collegano non soltanto al problema da cui sono nate, ma trovano applicazione in molti altri problemi. È proprio ciò che è avvenuto per la teoria degli insiemi. Le sue idee e i suoi concetti sono penetrati in tutti i rami della matematica, trasformandone profondamente l'aspetto. Non è quindi possibile formarsi un'immagine adeguata della matematica contemporanea senza conoscerne gli elementi. Per la teoria delle funzioni di una variabile reale, la teoria degli insiemi riveste un'importanza particolare.

Un insieme si considera assegnato quando è determinato, per ogni oggetto, se esso vi appartenga oppure no. In altri termini, un insieme è completamente determinato dagli oggetti che gli appartengono. Se un insieme M consiste degli oggetti a, b, c, ... e di nessun altro, scriviamo

$$M = \{a, b, c, \cdots\}.$$

Gli oggetti che formano un insieme si chiamano i suoi elementi. Il fatto che un oggetto m sia un elemento dell'insieme M si scrive

$$m \in M$$

e si legge: «m appartiene a M», oppure «m è un elemento di M». Se un oggetto n non appartiene a un insieme M, si scrive $n \bar{\in} M$. Ogni oggetto può comparire una sola volta come elemento di un dato insieme; in altri termini, gli elementi di un medesimo insieme sono tutti distinti.

Gli elementi di un insieme M possono essere a loro volta insiemi; per evitare contraddizioni si adotta qui, in particolare, la condizione che un insieme M non possa essere elemento di se stesso: $M \bar{\in} M$.

L'insieme che non contiene alcun elemento si chiama insieme vuoto. Per esempio, è vuoto l'insieme di tutte le radici reali dell'equazione

$$x^2 + 1 = 0$$

D'ora in avanti indicheremo l'insieme vuoto con ϕ.

Se ogni elemento x di un insieme M appartiene anche a un insieme N, diciamo che M è parte di N, che è un sottoinsieme di N, oppure che è contenuto in N; scriviamo

$$M \subseteq N \quad \text{oppure} \quad N \supseteq M.$$

Per esempio, l'insieme $M = \{1, 2\}$ è parte dell'insieme $N = \{1, 2, 3\}$.

È chiaro che vale sempre $M \subseteq M$. È conveniente considerare l'insieme vuoto come parte di ogni insieme.

Due insiemi sono uguali se hanno gli stessi elementi. Per esempio, l'insieme delle radici dell'equazione $x^2 - 3x + 2 = 0$ coincide con l'insieme $M = \{1, 2\}$.

Definiamo ora le operazioni sugli insiemi.

Unione o somma.

Siano M, N, P, ... degli insiemi. La loro unione, o somma, è l'insieme X degli elementi che appartengono ad almeno uno degli «addendi» M, N, P, ...

$$X = M + N + P + \cdots.$$

Anche se un elemento x appartiene a più addendi, nella somma X compare una volta sola. Evidentemente

$$M + M = M,$$

e, se $M \subseteq N$, allora

$$M + N = N.$$

Intersezione.

L'intersezione, o parte comune, degli insiemi M, N, P, ... è l'insieme Y degli elementi che appartengono a tutti gli insiemi M, N, P, ...

Evidentemente $M \cdot M = M$ e, se $M \subseteq N$, allora $M \cdot N = M$.

Se l'intersezione degli insiemi M e N è vuota, cioè $M \cdot N = \phi$, diciamo che i due insiemi sono disgiunti.

Per indicare l'unione e l'intersezione di insiemi useremo anche i simboli $\sum$ e $\prod$. Così,

$$E = \sum E_i$$

è la somma degli insiemi E_i, mentre

$$F = \prod E_i$$

è la loro intersezione.

Invitiamo il lettore a dimostrare che l'unione e l'intersezione sono legate dall'usuale legge distributiva

$$M(N + P) = MN + MP,$$

e anche dalla legge

$$M + NP = (M + N)(M + P).$$

Differenza.

La differenza di due insiemi M e N è l'insieme Z di tutti gli elementi di M che non appartengono a N:

$$Z = M - N.$$

Se $N \subseteq M$, la differenza $Z = M - N$ si chiama anche complementare di N in M.

Non è difficile dimostrare che vale sempre

$$M(N - P) = MN - MP$$

e

$$(M - N) + MN = M.$$

Le regole delle operazioni sugli insiemi differiscono dunque sensibilmente dalle consuete regole dell'aritmetica.

Insiemi finiti e infiniti.

Un insieme con un numero finito di elementi si chiama finito; un insieme che non è finito si chiama infinito. Per esempio, l'insieme di tutti i numeri naturali è infinito.

Consideriamo due insiemi arbitrari M e N e domandiamoci se abbiano oppure no lo stesso numero di elementi.

Se M è finito, la quantità dei suoi elementi è espressa da un numero intero non negativo, il numero degli elementi di M. Per confrontare due insiemi finiti basta dunque contare gli elementi di M e quelli di N e confrontare i numeri ottenuti. È inoltre naturale considerare che, se uno dei due insiemi è finito e l'altro infinito, quello infinito contenga più elementi.

Se però entrambi gli insiemi sono infiniti, il semplice conteggio non risolve il problema. Sorge quindi subito una domanda: tutti gli insiemi infiniti hanno lo stesso numero di elementi, oppure ne esistono con un numero maggiore o minore? E, in questo secondo caso, come confrontare il numero degli elementi di insiemi infiniti? Affronteremo ora questi problemi.

Corrispondenze biunivoche.

Siano ancora M e N due insiemi finiti. Come stabilire quale contenga più elementi senza contarli? Formiamo coppie associando a ogni elemento di M un elemento di N. Se, a un certo punto, rimane un elemento di M senza un elemento di N con cui accoppiarlo, M contiene più elementi. Illustriamo il ragionamento con un esempio.

In una stanza vi siano un certo numero di persone e un certo numero di sedie. Per stabilire quali siano più numerose basta invitare le persone a sedersi. Se qualcuno rimane in piedi, le persone sono più delle sedie; se tutti si siedono e tutte le sedie sono occupate, persone e sedie sono altrettante. Rispetto al conteggio diretto, questo metodo ha il vantaggio di applicarsi, senza modifiche sostanziali, anche agli insiemi infiniti.

Consideriamo l'insieme di tutti i numeri naturali

$$M = \{1, 2, 3, 4, \cdots\}$$

e l'insieme di tutti i numeri pari positivi

$$N = \{2, 4, 6, 8, \cdots\}.$$

Quale contiene più elementi? A prima vista sembrerebbe il primo. Possiamo tuttavia accoppiare gli elementi dei due insiemi come nella tavola 1.

Tavola 1.

M	1	2	3	4	$\cdots$
N	2	4	6	8	$\cdots$

Nessun elemento di M o di N rimane senza un corrispondente. Avremmo però potuto formare le coppie come nella tavola 2.

Tavola 2.

M	1	2	3	4	5	$\cdots$
N	—	2	—	4	—	$\cdots$

In questo caso molti elementi di M rimangono senza un corrispondente. Avremmo anche potuto accoppiarli come nella tavola 3.

Tavola 3.

M	—	1	—	2	—	3	—	⋯
N	2	4	6	8	10	12	14	⋯

Ora sono molti elementi di *N* a rimanere senza un corrispondente.

Dunque, se *A* e *B* sono infiniti, modi diversi di formare le coppie conducono a risultati diversi. Se esiste un modo in cui ogni elemento di *A* e ogni elemento di *B* appartengono esattamente a una coppia, diciamo che si può stabilire una corrispondenza biunivoca tra *A* e *B*. Per esempio, tra gli insiemi *M* e *N* si può stabilire una corrispondenza biunivoca, come mostra la tavola 1.

Se fra *A* e *B* si può stabilire una corrispondenza biunivoca, diciamo che hanno lo stesso numero di elementi. Se invece si possono accoppiare tutti gli elementi di *B* con elementi distinti di *A*, ma in ogni accoppiamento di questo tipo rimangono elementi di *A* senza corrispondente, diciamo che *A* contiene più elementi di *B*, oppure che ha cardinalità maggiore.

Abbiamo così risposto a una delle domande poste: come confrontare il numero degli elementi di insiemi infiniti. Non abbiamo però ancora risposto all'altra: esistono davvero insiemi infiniti di cardinalità diverse? Per rispondere, studiamo alcuni tipi semplici di insiemi infiniti.

Insiemi numerabili.

Se si può stabilire una corrispondenza biunivoca fra gli elementi di un insieme *A* e quelli dell'insieme di tutti i numeri naturali

$$Z = \{1, 2, 3, \cdots\},$$

diciamo che *A* è numerabile. In altre parole, un insieme è numerabile se tutti i suoi elementi si possono enumerare mediante i numeri naturali, cioè disporre in una successione

$$a_1 , a_2 , \cdots, a_n , \cdots.$$

La tavola 1 mostra che l'insieme dei numeri pari positivi è numerabile: i numeri della riga superiore possono essere interpretati come gli indici dei corrispondenti numeri della riga inferiore.

Gli insiemi numerabili sono, per così dire, i più piccoli insiemi infiniti: ogni insieme infinito contiene un sottoinsieme numerabile.

Se due insiemi finiti non vuoti sono disgiunti, la loro somma contiene più elementi di ciascun addendo. Per gli insiemi infiniti ciò non vale in generale. Sia, per esempio, *E* l'insieme dei numeri pari positivi, *O* quello dei numeri dispari positivi e *Z* quello dei numeri naturali. Come mostra la tavola 4, *E* e *O* sono numerabili. Eppure anche $Z = E + O$ è numerabile.

Tavola 4.

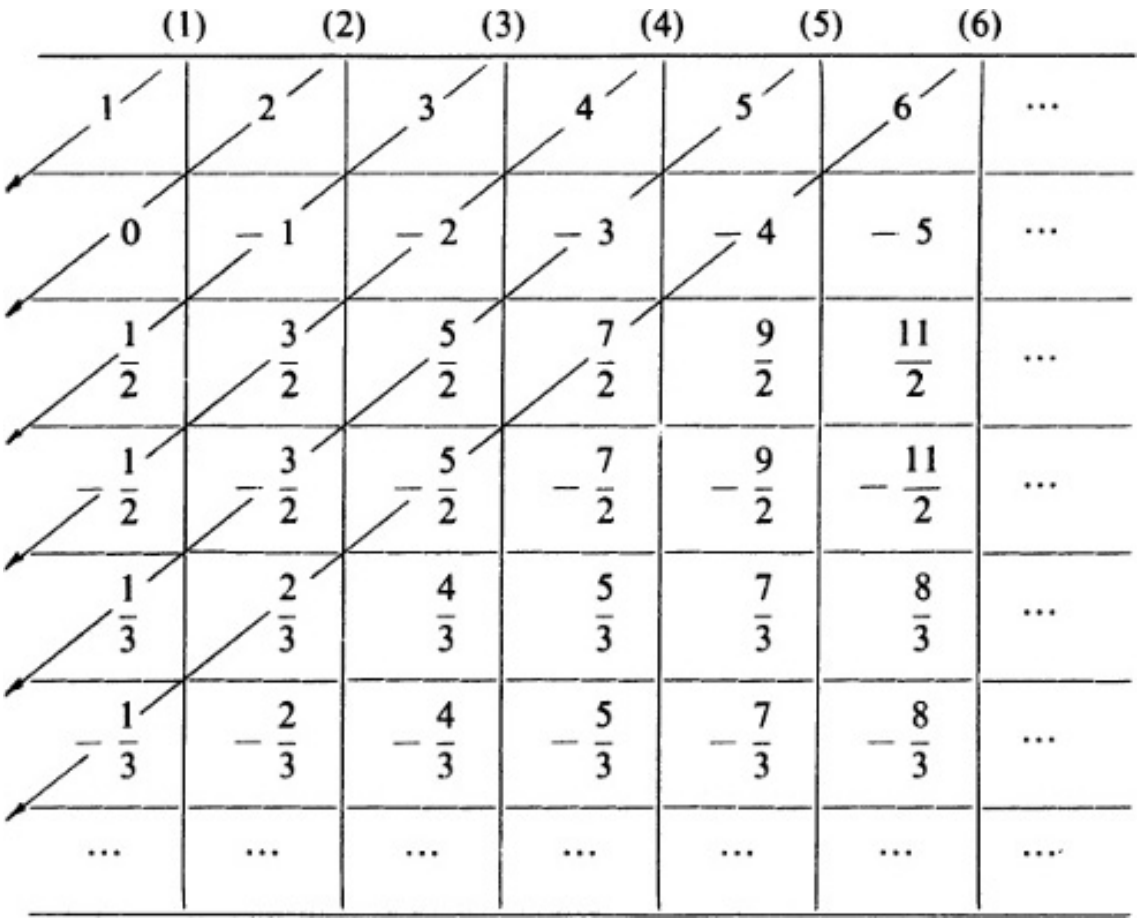
E	2	4	6	8	...
O	1	3	5	7	...
Z	1	2	3	4	...

Il venir meno della regola «il tutto è maggiore delle parti» mostra che le proprietà degli insiemi infiniti differiscono qualitativamente da quelle degli insiemi finiti. L'autore interpreta questo passaggio dal finito all'infinito come un esempio del principio dialettico della variazione qualitativa delle proprietà.

Dimostriamo che l'insieme di tutti i numeri razionali è numerabile. A questo scopo li disponiamo nella tavola 5.

Nella prima riga compaiono i numeri naturali in ordine crescente; nella seconda, lo zero e gli interi negativi in ordine decrescente; nella terza, le frazioni positive ridotte ai minimi termini con denominatore 2, in ordine crescente; nella quarta, le corrispondenti frazioni negative in ordine decrescente; e così via. Ogni numero razionale compare una e una sola volta nella tavola. Enumeriamo ora i numeri seguendo le frecce: tutti i razionali vengono così disposti in un'unica successione.

Tavola 5.



Posizione	1	2	3	4	5	6	7	8	9	...
Numero razionale	1	2	0	3	−1	1/2	4	−2	3/2	...

Abbiamo stabilito una corrispondenza biunivoca tra tutti i numeri razionali e tutti i numeri naturali. L'insieme dei numeri razionali è dunque numerabile.

Insiemi con la cardinalità del continuo.

Se si può stabilire una corrispondenza biunivoca fra gli elementi di un insieme *M* e i punti dell'intervallo $0 \leqslant x \leqslant 1$, diciamo che *M* ha la cardinalità del continuo. Per definizione, l'insieme dei punti del segmento $0 \leqslant x \leqslant 1$ ha esso stesso la cardinalità del continuo.

Dalla figura 1 è chiaro che l'insieme dei punti di qualsiasi segmento non degenere *AB* ha la cardinalità del continuo. La corrispondenza biunivoca si costruisce qui geometricamente mediante una proiezione.

Non è difficile mostrare che anche ogni intervallo aperto $a < x < b$, con $a < b$, e l'intera retta numerica hanno la cardinalità del continuo.

È assai più sorprendente il fatto seguente: l'insieme dei punti del quadrato $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant 1$ ha la cardinalità del continuo. In termini intuitivi, nel quadrato vi sono «tanti» punti quanti nel segmento.

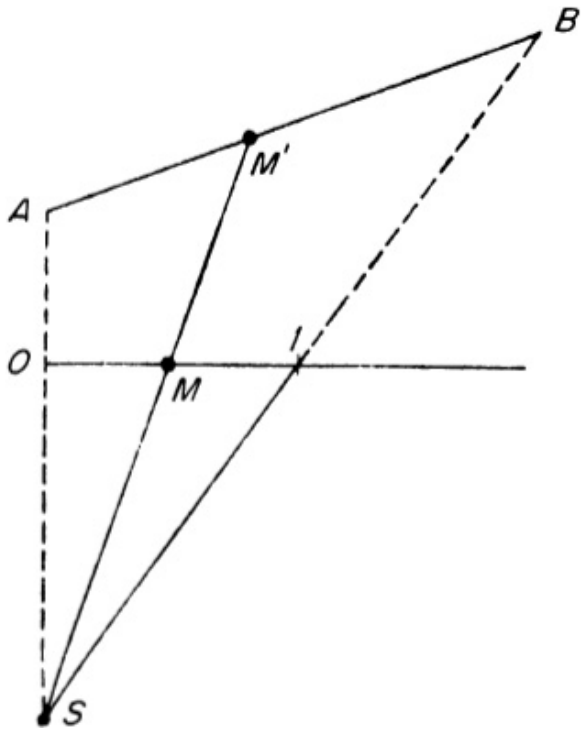


Figura 1. Corrispondenza biunivoca per proiezione fra due segmenti.

§3. Numeri reali[1]

Lo sviluppo del concetto di numero è stato descritto ampiamente nel capitolo I. Qui daremo un breve quadro delle teorie dei numeri reali sorte nel XIX secolo in relazione alla fondazione dei concetti fondamentali dell'analisi.

Numeri razionali.

Supponiamo note al lettore le principali proprietà dei numeri razionali e le richiamiamo senza entrare nei dettagli. I razionali, cioè i numeri della forma m/n, con m e n interi e $n \neq 0$, formano un insieme in cui sono definite due operazioni, addizione e moltiplicazione, soggette a determinate leggi, o assiomi. Nel seguito a, b, c indicano numeri razionali.

I. Assiomi dell'addizione.

1. $a + b = b + a$ (commutatività).

2. $a + (b + c) = (a + b) + c$ (associatività).

3. L'equazione

$$a + x = b$$

ha un'unica soluzione (esistenza dell'operazione inversa).

Da questi assiomi segue immediatamente che l'espressione $a + b + c$ ha un significato univoco, che esiste un numero razionale 0, detto zero, per cui $a + 0 = a$, e che l'addizione ammette un'operazione inversa, la sottrazione; l'espressione $b - a$ ha quindi significato.

Dal punto di vista algebrico, i numeri razionali formano dunque un gruppo commutativo rispetto all'addizione.

II. Assiomi della moltiplicazione.

1. $ab = ba$ (commutatività).

2. $a(bc) = (ab)c$ (associatività).

3. L'equazione

$$ay = b,$$

con $a \neq 0$ ha un'unica soluzione (esistenza dell'operazione inversa).

Da questi assiomi segue che l'espressione abc ha significato, che esiste un numero razionale 1 tale che $a \cdot 1 = a$ e che, per un divisore diverso da zero, esiste l'operazione inversa, la divisione. I razionali diversi da zero formano un gruppo commutativo rispetto alla moltiplicazione.

III. Assioma di distributività.

1. $(a + b)c = ac + bc$.

Gli assiomi I-III esprimono nel loro insieme il fatto che i numeri razionali, con le operazioni di addizione e moltiplicazione, formano un campo.

IV. Assiomi dell'ordine.

1. Per ogni coppia di razionali a e b vale una e una sola delle relazioni: $a < b$, $a > b$, $a = b$.

2. Se $a < b$ e $b < c$, allora $a < c$.

3. Se $a < b$, allora $a + c < b + c$ (monotonia dell'addizione).

4. Se $a < b$ e $c > 0$, allora $ac < bc$ (monotonia della moltiplicazione per $c > 0$).

Tutti questi assiomi permettono di chiamare l'insieme dei razionali un campo ordinato.

Esistono, oltre ai razionali, altri sistemi di oggetti che soddisfano gli stessi assiomi e costituiscono quindi campi ordinati.

Ricordiamo due proprietà importanti dei razionali.

Densità: dati a e b con $a < b$, esiste un razionale c tale che $a < c < b$.

Numerabilità: l'insieme di tutti i numeri razionali è numerabile (si veda il §2).

La misura delle grandezze.

L'insufficienza dei soli numeri razionali si manifesta già nel problema fondamentale della misura delle grandezze. Come esempio particolarmente semplice consideriamo la misura della lunghezza dei segmenti.

Fissiamo su una retta un verso positivo, un'origine, il punto 0, e un'unità di misura. È allora chiaro che cosa sia un segmento orientato OA il cui estremo abbia coordinata 1/2, 1/3, 2/3, −1/3, ecc. Più in generale, a ogni razionale a associamo il punto di coordinata $x = a$. Il numero a misura così il segmento orientato OA. Questa costruzione non assegna però a ogni segmento un numero razionale come misura della sua lunghezza. Già gli antichi Greci sapevano, per esempio, che la diagonale di un quadrato di lato unitario non può essere misurata da un numero razionale. È quindi naturale cercare una corrispondenza biunivoca fra numeri e misure dei segmenti orientati: occorre ampliare ancora il concetto di numero.

Numeri reali.

Siamo giunti alla conclusione che i soli razionali non bastano per misurare le grandezze e che il concetto di numero va esteso in modo da ottenere una corrispondenza biunivoca fra numeri e punti della retta. Cerchiamo allora di determinare la posizione di un punto arbitrario usando soltanto i punti razionali. Una costruzione analoga, svolta nel dominio dei numeri razionali, ci condurrà al concetto di numero reale.

Sia α un punto arbitrario della retta. Dividiamo i punti razionali a in due classi: quelli a sinistra di α e quelli a destra. Il punto α, se è razionale, può essere assegnato all'una o all'altra classe. Una simile divisione si chiama sezione. Consideriamo identiche due sezioni quando le rispettive classi coincidono, salvo l'eventuale assegnazione del singolo punto razionale di separazione. Punti distinti α e β determinano sezioni diverse: per la densità dei razionali, tra essi si trovano due punti razionali r_1 e r_2, che in una sezione appartengono alla classe di destra e nell'altra a quella di sinistra.

Ogni punto determina dunque una sezione dei punti razionali, e a punti diversi corrispondono sezioni diverse. È essenziale che la sezione si possa definire anche senza riferirsi a un punto α: è una divisione di tutti i punti razionali in due insiemi non vuoti e disgiunti A e B, tali che $a < b$ per ogni $a \in A$ e $b \in B$. A ogni sezione si associa univocamente il punto di separazione che la produce. Le sezioni dei punti razionali consentono così di determinare ogni punto della retta.

Questa costruzione, proposta dal matematico tedesco R. Dedekind, è nota come sezione di Dedekind.

La sezione non è l'unico metodo per individuare un punto arbitrario mediante punti razionali. Il procedimento di G. Cantor è più vicino alla pratica della misurazione. Sia ancora α un punto della retta. Possiamo trovare due punti razionali a e b, vicini quanto vogliamo, fra i quali sia compreso α. Essi ne determinano approssimativamente la posizione. Immaginiamo di proseguire indefinitamente, aumentando a ogni passo la precisione. Otteniamo intervalli $[a_n, b_n]$ con estremi razionali, tali che $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ e $b_n - a_n \to 0$ per $n \to \infty$. Un sistema con queste proprietà si chiama, qui, sistema di intervalli annidati. Esso determina univocamente la posizione di α.

Con costruzioni analoghe nel dominio dei razionali si definiscono i numeri reali. Si definiscono poi le loro operazioni e si verifica che soddisfano gli stessi assiomi delle operazioni sui razionali. A ogni punto della retta corrisponde ora un numero reale e viceversa; per questo l'insieme dei reali si chiama spesso retta numerica.

Principi di continuità.

Fra l'insieme dei razionali e quello dei reali esistono differenze essenziali. I reali possiedono proprietà che caratterizzano la loro continuità e che i razionali non possiedono. Queste proprietà prendono solitamente il nome di principi di continuità. Elenchiamo le principali.

Principio di Dedekind.

Se tutti i reali sono divisi in due insiemi non vuoti e disgiunti X e Y, tali che $x < y$ per ogni $x \in X$, $y \in Y$, esiste un unico numero ξ, detto elemento di separazione, per cui $x \leqslant \xi \leqslant y$ per ogni $x \in X$, $y \in Y$.

L'insieme dei reali x che soddisfano $a \leqslant x \leqslant b$ è un intervallo chiuso della retta numerica e si indica con $[a, b]$. Un sistema di intervalli $[a_n, b_n]$ è detto qui annidato se $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$ e $b_n - a_n \to 0$ per $n \to \infty$.

Principio di Cantor.

Per ogni sistema di intervalli annidati $[a_n, b_n]$ esiste uno e un solo reale ξ che appartiene a tutti gli intervalli.

Principio di Weierstrass.

Ogni successione non decrescente di numeri reali, limitata superiormente, converge.

Una successione di reali $\{x_n\}$ si chiama fondamentale, o di Cauchy, se per ogni $\varepsilon > 0$ esiste un naturale N tale che, per ogni $n > N$ e ogni naturale p,

$$| x_{n+p} - x_n | < \epsilon.$$

Principio di Cauchy.

Ogni successione fondamentale di numeri reali converge.

Non avendo costruito rigorosamente i numeri reali, non possiamo qui dimostrare che essi soddisfano questi principi. Studieremo invece i legami fra i principi stessi: supporremo che uno di essi valga e vedremo quali altri ne conseguano.

Il risultato complessivo sarà l'equivalenza dei quattro principi di continuità, con l'ipotesi archimedea precisata più avanti.

Un numero b si chiama estremo superiore di un insieme E, e si scrive

$$b = \sup E,$$

se: 1) $x \leqslant b$ per ogni $x \in E$; 2) non esiste un numero $b' < b$ con la stessa proprietà.

Dal principio di Dedekind segue che ogni insieme non vuoto E di numeri, limitato superiormente, ammette estremo superiore. Dividiamo tutti i reali in due classi X e Y: poniamo $x \in X$ se esiste $a \in E$ tale che $a \geqslant x$, e $y \in Y$ se $a < y$ per ogni $a \in E$. È facile verificare che si tratta di una sezione. Per il principio di Dedekind essa possiede un elemento di separazione ξ, che è l'estremo superiore di E.

Dimostriamo ora che dal principio di Dedekind segue quello di Weierstrass. Sia $\{x_n\}$ una successione non decrescente, limitata superiormente. Per quanto appena provato, l'insieme dei suoi termini ha estremo superiore ξ. Per definizione, $x_n \leqslant \xi$ $(n = 1, 2, \cdots)$; inoltre, per ogni $\varepsilon > 0$ esiste un indice n_0 tale che $x_{n_0} > \xi - \varepsilon$. Poiché la successione è monotona, ne segue $\xi - \epsilon < x_n \leqslant \xi$ per ogni $n > n_0$. La successione converge dunque a ξ.

Per dimostrare l'implicazione inversa osserviamo anzitutto che dal principio di Weierstrass segue il principio seguente.

Principio di Archimede.

Dati comunque due reali $a > 0$ e b, esiste un naturale n tale che $na > b$.

Questo principio equivale ad affermare che, per ogni reale b, la successione $\{b/n\}$, con $n = 1, 2, \ldots$, converge a zero.

Supponiamo valido il principio di Weierstrass e non quello di Archimede. Esisterebbe allora un $a > 0$ tale che la successione $x_n = na$ sia limitata. Essa è anche crescente, dunque per Weierstrass converge a un limite ξ. L'intervallo $[\xi - a/2, \xi]$ contiene perciò un termine $x_n = na$. Ma allora $x_{n+1} = (n + 1)a > \xi$, in contraddizione con il fatto che ξ è l'estremo superiore dei termini della successione.

Il principio di Weierstrass implica quello di Dedekind. Dividiamo tutti i reali in due insiemi non vuoti e disgiunti X e Y, con $x < y$ per ogni $x \in X$, $y \in Y$. Mostreremo che la sezione ha un unico elemento di separazione ξ. Fissato un naturale n, indichiamo con x_n il massimo numero della forma $m/2^n \in X$, con m intero; il suo successivo $x_n + 1/2^n$ appartiene a Y. L'esistenza di tale massimo segue dalla proprietà archimedea. I numeri della forma $m/2^n$ sono compresi fra quelli della forma $m/2^{n+1}$, dunque $x_n \leqslant x_{n+1}$. Inoltre $\{x_n\}$ è limitata superiormente, per esempio da $x_1 + 1/2$. Per Weierstrass converge a un limite ξ. Se $x < \xi$, allora $x \in X$. Se $y > \xi$, allora $y \in Y$: infatti, per Archimede possiamo scegliere n tale che $1/2^n < y - \xi = a$. Poiché $x_n \leq \xi$ e $x_n + 1/2^n \in Y$, risulta $y = \xi + a > x_n + 1/2^n$, quindi $y \in Y$. Questo prova che ξ è l'elemento di separazione, necessariamente unico.

Si può anche dimostrare l'equivalenza dei principi di Cantor e di Cauchy. Il principio di Cauchy, tuttavia, non implica da solo quello di Dedekind: esistono campi ordinati in cui vale il primo e non il secondo. Se si suppone fin dall'inizio valido il principio di Archimede, tutti e quattro i principi sono equivalenti.

Non numerabilità del continuo.

Dimostriamo che l'insieme dei punti del segmento $0 \leqslant x \leqslant 1$ non è numerabile. Ragioniamo per assurdo: supponiamo che l'insieme dei punti del segmento $0 \leqslant x \leqslant 1$ sia numerabile. Potremmo allora indicizzare tutti i suoi punti mediante i numeri naturali:

$$x_1 , x_2 , \cdots, x_n , \cdots. \qquad (1)$$

In $[0, 1]$ scegliamo un intervallo chiuso non degenere σ_1 di lunghezza minore di 1 che non contenga x_1. Ciò è sempre possibile. Dentro σ_1 scegliamo poi un intervallo chiuso non degenere σ_2, di lunghezza minore di $1/2$, che non contenga x_1 né x_2. In generale, scelto σ_{n-1}, scegliamo al suo interno un intervallo chiuso non degenere σ_n, di lunghezza minore di $1/n$, che non contenga $x_1, x_2, \ldots, x_n$. Otteniamo così una successione infinita di intervalli

$$\sigma_1 , \sigma_2 , \cdots, \sigma_n , \cdots,$$

ciascuno contenuto nel precedente, le cui lunghezze tendono a zero. Per il principio di Cantor esiste un unico punto x di $[0, 1]$ appartenente a tutti gli intervalli σ_n. Per ipotesi la successione (1) contiene tutti i punti di $[0, 1]$, dunque x coincide con un suo termine x_m. Ma per

costruzione σ_m non contiene x_m, mentre contiene x: una contraddizione. L'ipotesi iniziale è quindi falsa e l'insieme dei punti di [0, 1] non è numerabile, come volevamo dimostrare.

Questo teorema mostra che esistono cardinalità infinite diverse e risponde così affermativamente alla domanda posta.

§4. Insiemi di punti

Nella sezione precedente abbiamo già incontrato insiemi i cui elementi sono punti. In particolare, abbiamo considerato tutti i punti di un intervallo e tutti i punti (x, y) del quadrato $0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1$. Studieremo ora più da vicino le proprietà di questi insiemi.

Un insieme i cui elementi sono punti si chiama insieme di punti. Si parla così di insiemi di punti di una retta, di un piano o di uno spazio arbitrario. Per semplicità ci limiteremo agli insiemi di punti della retta.

Fra i numeri reali e i punti della retta esiste una stretta corrispondenza: a ogni reale associamo un punto e viceversa. Parlando di insiemi di punti comprendiamo perciò anche gli insiemi di numeri reali, cioè gli insiemi sulla retta numerica. Inversamente, per assegnare un insieme di punti della retta ne specificheremo, di regola, tutte le coordinate.

Gli insiemi di punti, e in particolare quelli della retta, hanno proprietà specifiche che li distinguono dagli insiemi arbitrari e fanno della loro teoria una disciplina matematica autonoma. Anzitutto ha senso parlare di distanza fra due punti. Sulla retta si può inoltre introdurre un ordine, distinguendo destra e sinistra; un insieme di punti della retta è quindi un insieme ordinato. Infine sulla retta vale il principio di Cantor, proprietà che esprime la sua completezza.

Introduciamo alcune notazioni per gli insiemi più semplici della retta.

L'intervallo chiuso $[a, b]$ è l'insieme dei punti le cui coordinate soddisfano $a \leqslant x \leqslant b$.

L'intervallo aperto (a, b) è l'insieme dei punti le cui coordinate soddisfano $a < x < b$.

Gli intervalli semiaperti $[a, b)$ e $(a, b]$ sono definiti rispettivamente dalle condizioni $a \leqslant x < b$ e $a < x \leqslant b$.

Gli intervalli possono anche essere illimitati. Così $(-\infty, \infty)$ indica l'intera retta e, per esempio, $(-\infty, b]$ l'insieme dei punti per cui $x \leqslant b$.

Cominciamo dalle diverse possibilità di collocazione di un insieme nel suo complesso sulla retta.

Insiemi limitati e illimitati.

Un insieme E della retta può essere formato da punti la cui distanza dall'origine non supera un certo numero positivo, oppure contenere punti arbitrariamente lontani dall'origine. Nel primo caso si dice limitato, nel secondo illimitato. L'insieme dei punti di [0, 1] è limitato; quello dei punti con coordinate intere è illimitato.

Fissato un punto a della retta, è facile vedere che E è limitato se e solo se le distanze da a di tutti i punti $x \in E$ non superano un certo numero positivo.

Insiemi limitati superiormente e inferiormente.

Un insieme E della retta si dice limitato superiormente se esiste un punto A tale che ogni $x \in E$ si trovi alla sua sinistra. Analogamente, se esiste un punto a tale che ogni $x \in E$ si trovi alla sua destra, E si dice limitato inferiormente. L'insieme dei punti con coordinate positive è limitato inferiormente, quello dei punti con coordinate negative è limitato superiormente.

La definizione di insieme limitato equivale dunque a questa: un insieme E della retta è limitato se è limitato sia superiormente sia inferiormente. Nonostante la loro somiglianza, le due definizioni si basano su aspetti differenti: la prima sulla distanza fra i punti, la seconda sull'ordine della retta.

Possiamo anche dire che un insieme è limitato se è interamente contenuto in qualche intervallo $[a, b]$.

Estremo superiore ed estremo inferiore.

Sia E un insieme non vuoto limitato superiormente. Esistono allora punti A alla cui destra non si trova alcun punto di E. Servendosi del principio di Cantor si dimostra che fra tutti questi punti A ve n'è uno più a sinistra di tutti gli altri: è l'estremo superiore di E. Analogamente si definisce l'estremo inferiore.

Se E possiede un punto più a destra di tutti gli altri, questo è evidentemente il suo estremo superiore. Ma tale punto può non esistere.

Per esempio, l'insieme dei punti di coordinate

$$\frac{0}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \cdots$$

è limitato superiormente, ma non ha un punto più a destra di tutti gli altri. Il suo estremo superiore a non appartiene a E, mentre esistono punti di E arbitrariamente vicini ad a. Nell'esempio, $a = 1$.

Distribuzione di un insieme presso un punto della retta.

Sia *E* un insieme di punti e *x* un punto arbitrario della retta. Consideriamo come può disporsi *E* nelle vicinanze di *x*. Sono possibili quattro casi.

1. Né *x* né i punti sufficientemente vicini a esso appartengono a *E*.

2. Il punto *x* non appartiene a *E*, ma esistono punti di *E* arbitrariamente vicini a esso.

3. Il punto *x* appartiene a *E*, ma tutti gli altri punti sufficientemente vicini a esso non vi appartengono.

4. Il punto *x* appartiene a *E* ed esistono altri punti di *E* arbitrariamente vicini a esso.

Nel caso 1, *x* si dice esterno a *E*; nel caso 3, punto isolato di *E*; nei casi 2 e 4, punto di accumulazione di *E*.

Se dunque $x \bar{\in} E$, *x* può essere esterno oppure di accumulazione; se $x \in E$, può essere isolato oppure di accumulazione.

Un punto di accumulazione può appartenere o non appartenere a *E*; ciò che lo caratterizza è l'esistenza di punti di *E*, diversi da esso, arbitrariamente vicini. Equivalentemente, *x* è di accumulazione per *E* se ogni intervallo aperto δ che contiene *x* contiene infiniti punti di *E*. Questo è uno dei concetti più importanti della teoria degli insiemi di punti.

Se *x* e tutti i punti sufficientemente vicini a esso appartengono a *E*, *x* si dice interno a *E*. Un punto che non è né esterno né interno si chiama punto di frontiera di *E*.

Illustriamo questi concetti con alcuni esempi.

Esempio 1. Sia E_1 l'insieme dei punti di coordinate

$$1, \frac{1}{2}, \frac{1}{3}, \cdots, \frac{1}{n}, \cdots .$$

Ogni punto dell'insieme è isolato. Il punto 0 è di accumulazione per E_1, ma non vi appartiene; tutti gli altri punti della retta sono esterni a E_1.

Esempio 2. Sia E_2 l'insieme dei punti razionali di [0, 1]. Non ha punti isolati; ogni punto di [0, 1] è di accumulazione per E_2, mentre tutti gli altri punti della retta sono esterni. Fra i punti di accumulazione, alcuni appartengono all'insieme e altri no.

Esempio 3. Sia E_3 l'insieme di tutti i punti di [0, 1]. Come nell'esempio precedente, non ha punti isolati e ogni punto di [0, 1] è di accumulazione. Questa volta, però, tutti i punti di accumulazione appartengono all'insieme.

Esempio 4. Sia E_4 l'insieme dei punti della retta con coordinate intere. Ogni suo punto è isolato e E_4 non ha punti di accumulazione.

Osserviamo inoltre che nell'esempio 3 ogni punto dell'intervallo aperto (0, 1) è interno a E_3, mentre nell'esempio 2 ogni punto di [0, 1] è di frontiera per E_2.

Un insieme infinito della retta può dunque avere punti isolati, come E_1 ed E_4, oppure non averne, come E_2 ed E_3; può avere punti interni, come E_3, oppure non averne, come E_1, E_2, E_4. Solo E_4 non ha punti di accumulazione. Il teorema seguente mostra che ciò è legato al fatto che esso è illimitato.

Teorema di Bolzano-Weierstrass.

Ogni insieme infinito e limitato di punti della retta possiede almeno un punto di accumulazione.

Dimostrazione. Sia E un insieme infinito e limitato della retta. Essendo limitato, è contenuto in un intervallo $[a, b]$. Dividiamo l'intervallo in due metà chiuse: almeno una contiene infiniti punti di E. Chiamiamola σ_1; se entrambe li contengono, scegliamo per esempio quella di sinistra. Dividiamo σ_1 in due metà uguali e scegliamo una metà σ_2 contenente ancora infiniti punti di E. Proseguendo indefinitamente otteniamo gli intervalli σ_1, σ_2, ..., σ_n, ..., con tre proprietà: ciascuno è contenuto nel precedente; ciascuno contiene infiniti punti di E; le loro lunghezze tendono a zero. Le prime due proprietà seguono dalla costruzione; per la terza basta osservare che, se $[a, b]$ ha lunghezza l, allora σ_n ha lunghezza $l/2^n$. Per il principio di Cantor esiste un unico punto x comune a tutti gli intervalli. Dimostriamo che è di accumulazione per E. Sia δ un intervallo aperto contenente x. Poiché tutti gli intervalli σ_n contengono x e hanno lunghezza tendente a zero, per n sufficientemente grande σ_n è interamente contenuto in δ. Ma σ_n contiene infiniti punti di E, dunque anche δ li contiene. Il punto x è perciò di accumulazione e il teorema è dimostrato.

Esercizio. Dimostrare che, se un insieme non vuoto E è limitato superiormente e non ha massimo, il suo estremo superiore è un punto di accumulazione di E e non appartiene a E.

Insiemi chiusi e aperti.

Uno dei compiti fondamentali della teoria degli insiemi di punti è lo studio delle proprietà dei loro diversi tipi. Ne illustreremo due particolarmente importanti: gli insiemi chiusi e gli insiemi aperti.

Un insieme si dice chiuso se contiene tutti i suoi punti di accumulazione. Anche un insieme senza punti di accumulazione è chiuso. Oltre ai punti di accumulazione, un chiuso può contenere punti isolati. Un insieme si dice aperto se ogni suo punto è interno.

Per esempio, $[a, b]$ è chiuso e (a, b) è aperto. Le semirette $(-\infty, b]$ e $[a, \infty)$ sono chiuse, mentre $(-\infty, b)$ e (a, ∞) sono aperte. L'intera retta è insieme aperta e chiusa; lo stesso vale per l'insieme vuoto. Ogni insieme finito della retta è chiuso, perché non ha punti di accumulazione. Anche l'insieme dei punti

$$0, 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots \frac{1}{n}, \cdots$$

è chiuso: il suo unico punto di accumulazione è $x = 0$, che appartiene all'insieme.

Vogliamo esaminare la struttura di un insieme chiuso o aperto arbitrario. Ci serviremo delle seguenti proprietà ausiliarie, che ammettiamo senza dimostrazione.

1. L'intersezione di una famiglia arbitraria di insiemi chiusi è chiusa.

2. L'unione di una famiglia arbitraria di insiemi aperti è aperta.

3. Un insieme chiuso non vuoto, limitato superiormente, contiene il proprio estremo superiore. Analogamente, se è limitato inferiormente, contiene il proprio estremo inferiore.

Sia E un insieme arbitrario della retta. Il suo complementare, indicato con CE, è l'insieme di tutti i punti della retta che non appartengono a E. Evidentemente un punto è esterno a E se e solo se è interno a CE.

4. Un insieme F è chiuso se e solo se il suo complementare CF è aperto.

La proprietà 4 rivela il legame strettissimo fra aperti e chiusi: sono i complementari gli uni degli altri. Basta quindi studiare uno dei due tipi; conoscendone le proprietà si ricavano immediatamente quelle dell'altro. Per esempio, ogni aperto si ottiene togliendo dalla retta un insieme chiuso.

Studiamo ora i chiusi. Sia F un insieme chiuso. Un intervallo aperto (a, b) che non contiene punti di F, ma i cui estremi a e b appartengono a F, si chiama intervallo contiguo a F. Comprendiamo fra gli intervalli contigui anche (a, ∞) e $(-\infty, b)$, purché l'estremo finito appartenga a F e la semiretta non lo incontri. Dimostreremo che ogni punto x esterno a un chiuso non vuoto F appartiene a uno dei suoi intervalli contigui.

Indichiamo con F_x la parte di F situata a destra di x. Poiché x non appartiene a F, possiamo scrivere

$$F_x = F \cdot [x, \infty).$$

Sia F sia $[x, \infty)$ sono chiusi; per la proprietà 1 lo è anche F_x. Se F_x è vuoto, l'intera semiretta $[x, \infty)$ non incontra F. Se non è vuoto, è limitato inferiormente; indichiamo con b il suo estremo inferiore. Per la proprietà 3, $b \in F_x$, quindi $b \in F$. L'intervallo $[x, b)$ non contiene punti di F_x, né dunque di F. Abbiamo così costruito $[x, b)$ senza punti di F, con $b = \infty$ oppure $b \in F$. Analogamente si costruisce $(a, x]$ senza punti di F, con $a = -\infty$ oppure $a \in F$. L'intervallo (a, b) contiene x ed è contiguo a F. Si vede facilmente che due intervalli contigui (a_1, b_1) e (a_2, b_2) o coincidono o sono disgiunti.

Ogni chiuso della retta si ottiene quindi eliminando dalla retta i suoi intervalli contigui. Ogni intervallo aperto contiene almeno un punto razionale; scegliendone uno in ciascun intervallo contiguo, si ottengono punti distinti. Poiché i razionali formano un insieme numerabile, gli intervalli contigui sono al più numerabili. Giungiamo così a una conclusione notevole: ogni chiuso della retta si ottiene eliminando dalla retta una famiglia al più numerabile di intervalli aperti disgiunti. Il caso del chiuso vuoto si ottiene eliminando l'intera retta, intesa come intervallo aperto illimitato.

Dalla proprietà 4 segue subito che ogni aperto della retta è unione di una famiglia al più numerabile di intervalli aperti disgiunti. Le proprietà 1 e 2 mostrano anche che gli insiemi costruiti in questo modo sono effettivamente, rispettivamente, chiusi e aperti.

L'esempio seguente mostra quanto possa essere complessa la struttura di un insieme chiuso.

L'insieme perfetto di Cantor.

Costruiamo un particolare chiuso con proprietà notevoli. Togliamo anzitutto dalla retta gli intervalli $(-\infty, 0)$ e $(1, \infty)$: rimane $[0, 1]$. Togliamo poi il suo terzo centrale aperto, $(1/3, 2/3)$. Da ciascuno dei segmenti rimasti, $[0, 1/3]$ e $[2/3, 1]$, togliamo di nuovo il terzo centrale aperto. Ripetiamo indefinitamente il procedimento. L'insieme che rimane dopo tutte queste eliminazioni è l'insieme perfetto di Cantor; lo indicheremo con P.

Esaminiamone alcune proprietà. P è chiuso, perché si ottiene eliminando dalla retta un insieme di intervalli aperti disgiunti. Non è vuoto: contiene almeno tutti gli estremi finiti degli intervalli eliminati.

Un chiuso F si chiama perfetto se non ha punti isolati, cioè se ogni suo punto è di accumulazione. Mostriamo che P è perfetto. Se x fosse un suo punto isolato, dovrebbe essere estremo comune di due intervalli contigui. Per costruzione, però, gli intervalli contigui a P non hanno estremi comuni.

P non contiene alcun intervallo aperto non vuoto. Infatti, se un intervallo aperto δ fosse contenuto in P, sarebbe contenuto in uno dei segmenti rimasti al passo n della costruzione, per ogni n. Ma ciò è impossibile, perché le lunghezze di quei segmenti tendono a zero.

Si può dimostrare che P ha la cardinalità del continuo. In particolare, contiene punti diversi dagli estremi degli intervalli contigui, poiché questi ultimi formano soltanto un insieme numerabile.

Insiemi di punti di vario tipo compaiono continuamente nei rami più diversi della matematica; conoscerne le proprietà è indispensabile per molti problemi. La loro teoria ha particolare importanza nell'analisi matematica e nella topologia.

Consideriamo alcuni esempi che sorgono nell'analisi classica. Sia $f(x)$ continua su $[a, b]$. Fissiamo α e consideriamo l'insieme dei punti per cui $f(x) \geqslant \alpha$: esso è chiuso e, scegliendo opportunamente la funzione, può essere un qualsiasi chiuso di $[a, b]$. Analogamente, i punti per cui $f(x) > \alpha$ formano un aperto relativo di $[a, b]$, e ogni aperto relativo può essere ottenuto così. Se invece $f_1(x), f_2(x), \ldots$ è una successione di funzioni continue su $[a, b]$, l'insieme dei punti in cui essa converge non è arbitrario: appartiene a una classe ben determinata.

La disciplina che studia la struttura degli insiemi di punti si chiama teoria descrittiva degli insiemi.

Al suo sviluppo hanno contribuito in modo importante i matematici sovietici N. N. Luzin e i suoi allievi, P. S. Aleksandrov, M. Ja. Suslin, A. N. Kolmogorov, M. A. Lavrent'ev, P. S. Novikov, L. V. Keldyš, A. A. Ljapunov e altri.

Le ricerche di Luzin e dei suoi allievi hanno mostrato gli stretti rapporti fra teoria descrittiva degli insiemi e logica matematica. Le difficoltà di alcuni problemi, in particolare quelli sulla cardinalità di certi insiemi, hanno natura logica. D'altra parte, i metodi della logica matematica consentono di approfondire alcuni problemi della teoria descrittiva.

§5. Misura degli insiemi

Il concetto di misura di un insieme generalizza ampiamente quello di lunghezza di un intervallo. Nel caso più semplice, l'unico che considereremo, si tratta di definire una lunghezza non soltanto per gli intervalli, ma anche per insiemi più complessi della retta.

Assumiamo come unità di misura [0, 1]. La lunghezza di $[a, b]$ è allora $b - a$. Se due intervalli $[a_1, b_1]$ e $[a_2, b_2]$ sono disgiunti, è naturale attribuire alla loro unione la lunghezza $(b_1 - a_1) + (b_2 - a_2)$. Non è invece evidente che cosa si debba intendere per lunghezza di un insieme più complesso: per esempio, qual è la lunghezza dell'insieme di Cantor del §4? Occorre dunque una definizione matematica rigorosa.

Definire la lunghezza, o, come diremo ora, la misura degli insiemi è fondamentale per generalizzare il concetto di integrale. La misura trova inoltre applicazioni in altri problemi della teoria delle funzioni, nella probabilità, nella topologia, nell'analisi funzionale e altrove.

Esporremo la definizione dovuta al matematico francese H. Lebesgue, sulla quale si fonda la sua definizione di integrale.

Misura di un aperto e di un chiuso.

Cominciamo dagli insiemi aperti e chiusi. Come abbiamo visto nel §4, ogni aperto della retta è unione finita o numerabile di intervalli aperti a due a due disgiunti.

La misura di un aperto è, per definizione, la somma delle lunghezze degli intervalli aperti che lo compongono.

Se dunque

$$G = \sum (a_i , b_i)$$

e gli intervalli (a_i, b_i) sono a due a due disgiunti, la misura di G è $\sum (b_i - a_i)$. Indicando in generale con μE la misura di un insieme E, scriviamo

$$\mu G = \sum (b_i - a_i).$$

In particolare, la misura di un solo intervallo aperto è la sua lunghezza:

$$\mu(a, b) = b - a.$$

Un chiuso F contenuto in $[a, b]$ e comprendente entrambi gli estremi si ottiene togliendo da $[a, b]$ un aperto G. La sua misura si definisce come la differenza fra la lunghezza di $[a, b]$ e la misura di G, complementare di F rispetto a $[a, b]$.

Dunque

$$\mu F = (b - a) - \mu G. \qquad (2)$$

È facile verificare che, secondo questa definizione, la misura di un intervallo chiuso coincide con la sua lunghezza:

$$\mu[a, b] = b - a,$$

mentre la misura di un insieme finito di punti è zero.

Definizione generale di misura.

Per passare a insiemi più generali degli aperti e dei chiusi occorre un concetto ausiliario. Sia E contenuto in $[a, b]$. Consideriamo tutti i suoi ricoprimenti aperti, cioè tutti gli aperti $V(E)$ che lo contengono. La loro misura è già definita. Le misure dei ricoprimenti aperti limitati formano un insieme non vuoto di numeri non negativi, limitato inferiormente da zero; esso ha quindi un estremo inferiore, che indichiamo con $\mu_e E$. Questo numero si chiama misura esterna di E.

Siano $\mu_e E$ la misura esterna di E e $\mu_e CE$ quella del suo complementare rispetto a $[a, b]$.

Se vale la relazione

$$\mu_e E + \mu_e CE = b - a \qquad (3)$$

l'insieme E si dice misurabile e il numero $\mu_e E$ è la sua misura: $\mu E = \mu_e E$. Se la relazione (3) non vale, E si dice non misurabile e non gli viene attribuita una misura di Lebesgue.

Osserviamo che vale sempre

$$\mu_e E + \mu_e CE \geqslant b - a. \qquad (4)$$

Aggiungiamo qualche chiarimento. La lunghezza degli insiemi più semplici, per esempio degli intervalli aperti o chiusi, ha proprietà importanti. Elenchiamo quelle fondamentali che la misura deve conservare.

1. Se E_1 ed E sono misurabili e $E_1 \subseteq E$, allora

$$\mu E_1 \leqslant \mu E;$$

cioè la misura di una parte non supera quella dell'insieme intero.

2. Se E_1 ed E_2 sono misurabili, anche $E = E_1 + E_2$ è misurabile e

$$\mu(E_1 + E_2) \leqslant \mu E_1 + \mu E_2;$$

cioè la misura di un'unione non supera la somma delle misure degli insiemi che la compongono.

3. Se gli insiemi E_i, con $i = 1, 2, \ldots$, sono misurabili e a due a due disgiunti, $E_iE_j = \phi$ per $i \neq j$, la loro somma $E = \sum E_i$ è misurabile e

$$\mu(\sum E_i) = \sum \mu E_i ;$$

cioè la misura di un'unione finita o numerabile di insiemi disgiunti è la somma delle loro misure.

Questa proprietà si chiama additività completa, o numerabile, della misura.

4. La misura di un insieme E non cambia se lo si sposta rigidamente.

Si desidera che le proprietà fondamentali della lunghezza rimangano valide per il concetto più generale di misura. Si può però dimostrare rigorosamente che, nell'usuale teoria degli insiemi, è impossibile conservarle tutte attribuendo una misura a ogni insieme di punti della retta. Esistono dunque insiemi misurabili e insiemi non misurabili. La classe dei misurabili è tuttavia tanto ampia che questa limitazione non comporta inconvenienti sostanziali nelle applicazioni. Costruire un esempio di insieme non misurabile richiede, anzi, un procedimento non immediato.

Presentiamo alcuni esempi di insiemi misurabili.

Esempio 1: misura dell'insieme perfetto di Cantor. Nella costruzione di P a partire da [0, 1] abbiamo eliminato dapprima un intervallo di lunghezza 1/3, poi due intervalli di lunghezza 1/9, poi quattro di lunghezza 1/27, e così via. Al passo n abbiamo eliminato 2^{n-1} intervalli, ciascuno di lunghezza $1/3^n$. La somma delle lunghezze eliminate è quindi

$$S = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots + \frac{2^{n-1}}{3^n} + \cdots .$$

I termini formano una progressione geometrica di primo termine 1/3 e ragione 2/3. La somma S della serie è perciò

$$\frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1.$$

La somma delle lunghezze degli intervalli eliminati da [0, 1] è dunque 1: l'aperto *G*, complementare di *P* rispetto a [0, 1], ha misura 1. Ne segue che

$$\mu P = 1 - \mu G = 1 - 1 = 0.$$

Questo esempio mostra che un insieme può avere la cardinalità del continuo e tuttavia misura zero.

Esempio 2: misura dei punti razionali di [0, 1]. Indichiamo questo insieme con *R* e dimostriamo anzitutto che $\mu_e R = 0$. Nel §2 abbiamo visto che *R* è numerabile. Disponiamone i punti in successione:

$$r_1 , r_2 , \cdots, r_n , \cdots .$$

Fissato $\varepsilon > 0$, racchiudiamo il punto r_n in un intervallo aperto δ_n di lunghezza $\varepsilon/2^n$. L'unione $G = \sum \delta_n$ è un aperto che ricopre *R*. Gli intervalli possono sovrapporsi, perciò

$$\mu G = \mu(\sum \delta_n) \leq \sum \mu\delta_n = \sum \varepsilon/2^n = \varepsilon.$$

Poiché ε può essere scelto piccolo quanto si vuole, $\mu_e R = 0$.

Inoltre, per la disuguaglianza generale fra le misure esterne di un insieme e del suo complementare,

$$\mu_e R + \mu_e CR \geqslant 1,$$

cioè $\mu_e CR \geqslant 1$. D'altra parte *CR* è contenuto in [0, 1], quindi $\mu_e CR \leqslant 1$.

Pertanto

$$\mu_e R + \mu_e CR = 1,$$

e[2]

$$\mu R = 0, \quad \mu CR = 1. \qquad (5)$$

Questo esempio mostra che un insieme può essere denso in tutto un intervallo e avere tuttavia misura zero.

In molti problemi della teoria delle funzioni gli insiemi di misura zero possono essere trascurati. Per esempio, una funzione su un intervallo chiuso e limitato è integrabile secondo Riemann se e solo se è limitata e l'insieme dei suoi punti di discontinuità ha misura zero. Si potrebbero aggiungere molti altri esempi.

Funzioni misurabili.

Passiamo a una delle applicazioni più feconde della misura: la descrizione della classe di funzioni con cui operano l'analisi e la teoria delle funzioni. Precisiamo il problema. Se una successione $\{f_n(x)\}$ di funzioni definite su un insieme E converge in ogni punto di E, eccetto eventualmente quelli di un insieme N di misura zero, diciamo che converge quasi ovunque.

Quali funzioni si possono ottenere dalle funzioni continue applicando ripetutamente le operazioni algebriche e il passaggio al limite di successioni convergenti quasi ovunque, con libertà di assegnare i valori sull'insieme eccezionale di misura zero?

Per rispondere occorrono alcuni nuovi concetti.

Sia $f(x)$ una funzione definita su E e sia α un reale arbitrario. Indichiamo con

$$E[f(x) > \alpha]$$

l'insieme dei punti di E per cui $f(x) > \alpha$. Per esempio, se $f(x) = x$ su [0, 1], tale insieme coincide con [0, 1] per $\alpha < 0$, con $(\alpha, 1]$ per $0 \leqslant \alpha < 1$, ed è vuoto per $\alpha \geqslant 1$.

Una funzione $f(x)$ definita su E si dice misurabile se E è misurabile e, per ogni reale α, è misurabile l'insieme $E[f(x) > \alpha]$.

Ogni funzione continua su un intervallo è misurabile. Ma sono misurabili anche molte funzioni discontinue: per esempio la funzione di Dirichlet, nella convenzione qui adottata, uguale a 1 nei punti irrazionali di [0, 1] e a 0 nei punti razionali.

Enunciamo senza dimostrazione alcune proprietà.

1. Se $f(x)$ e $\phi(x)$ sono misurabili sullo stesso insieme E, lo sono anche le funzioni

$$f + \phi,\ f - \phi,\ f \cdot \phi \quad \text{e} \quad f/\phi$$

l'ultima a condizione che ϕ non si annulli.

Le operazioni algebriche sulle funzioni misurabili producono dunque ancora funzioni misurabili.

2. Se una successione $\{f_n(x)\}$ di funzioni misurabili su E converge quasi ovunque a una funzione $f(x)$, anche questa è misurabile.

Anche il passaggio al limite di una successione convergente quasi ovunque conserva dunque la misurabilità.

Queste proprietà furono stabilite da Lebesgue. Lo studio delle funzioni misurabili fu approfondito dai matematici sovietici D. F. Egorov e N. N. Luzin. In particolare, Luzin dimostrò che una funzione reale misurabile su un intervallo chiuso e limitato può essere resa continua modificandone i valori su un insieme di misura arbitrariamente piccola.

Questo risultato classico di Luzin e le proprietà appena elencate permettono di riconoscere nelle funzioni misurabili la classe descritta all'inizio della sezione. Esse hanno anche un ruolo fondamentale nella teoria dell'integrazione: il concetto di integrale può essere esteso in modo che ogni funzione misurabile e limitata su un intervallo limitato sia integrabile. Lo vedremo nella prossima sezione.

§6. L'integrale di Lebesgue

Affrontiamo ora il tema centrale del capitolo: la definizione dell'integrale di Lebesgue e le sue proprietà.

Per comprenderne l'idea, immaginiamo una grande raccolta di monete di tagli diversi e supponiamo di volerne calcolare il valore totale. Possiamo procedere in due modi. Disponiamo le monete in fila e aggiungiamo il valore di ciascuna alla somma di quelle precedenti. Oppure raggruppiamo le monete per taglio, contiamo quante ve ne sono in ciascun gruppo, moltiplichiamo il numero per il valore del taglio e sommiamo i risultati. Il primo metodo corrisponde all'integrazione di Riemann, il secondo a quella di Lebesgue.

Passando dalle monete alle funzioni, per calcolare l'integrale di Riemann si suddivide in piccole parti il dominio della funzione, sull'asse delle ascisse (figura 2a); per quello di Lebesgue si suddivide invece l'intervallo dei valori, sull'asse delle ordinate (figura 2b). Questo secondo principio era usato nella pratica già prima di Lebesgue per integrare funzioni oscillanti. Lebesgue fu però il primo a svilupparlo in piena generalità e a fondarlo rigorosamente sulla teoria della misura.

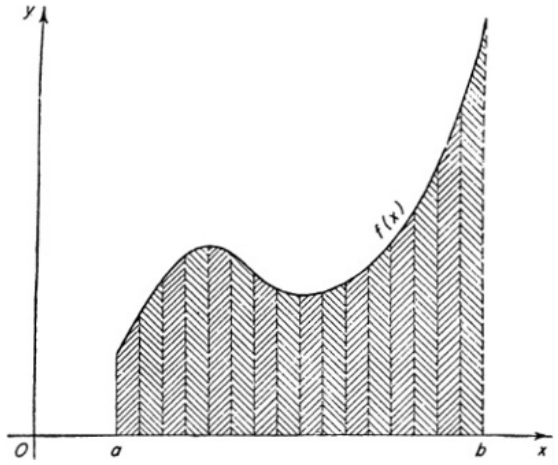


Figura 2a. Integrazione di Riemann: suddivisione dell'intervallo delle ascisse.

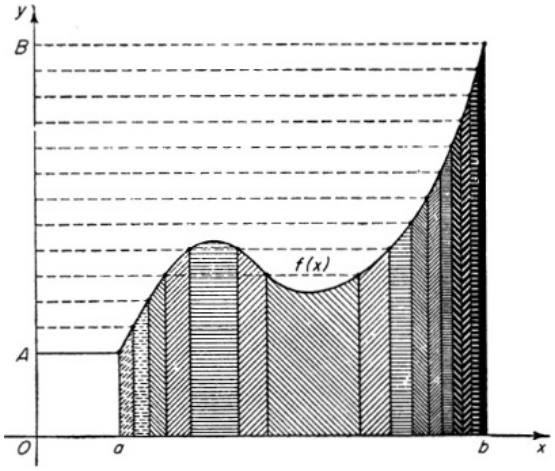


Figura 2b. Integrazione di Lebesgue: suddivisione dell'intervallo dei valori.

Vediamo il legame fra misura e integrale. Sia E un insieme misurabile contenuto in $[a, b]$. Costruiamo la funzione $\phi(x)$ uguale a 1 se x appartiene a E e a zero altrimenti. Poniamo cioè

$$\phi(x) = \begin{cases} 1 & x \in E. \\ 0 & x \bar{\in} E. \end{cases}$$

La funzione $\phi(x)$ si chiama funzione caratteristica di E. Consideriamo l'integrale

$$I = \int_a^b \phi(x)\, dx.$$

Siamo abituati a interpretare l'integrale come l'area della figura D compresa fra l'asse delle ascisse, le rette $x = a$, $x = b$ e il grafico $y = \phi(x)$ (capitolo II). Nel nostro caso l'«altezza» di D è 1 esattamente sopra i punti di E ed è zero altrove. Secondo l'idea «area uguale base per altezza», l'area dovrebbe quindi coincidere numericamente con la lunghezza, cioè la misura, di E. Vorremo dunque avere

$$I = \mu E. \qquad (6)$$

L'integrale di Lebesgue di $\phi(x)$ viene definito proprio così.

È essenziale comprendere che l'uguaglianza (6) costituisce la definizione di $\int_a^b \phi(x)\, dx$ come integrale di Lebesgue. L'integrale I può non esistere nel senso del capitolo II, cioè come limite delle somme di Riemann; anche in quel caso, però, esiste come integrale di Lebesgue ed è uguale a μE.

Calcoliamo, per esempio, l'integrale della funzione di Dirichlet $\Phi(x)$, uguale a 0 nei punti razionali e a 1 negli irrazionali di [0, 1]. Poiché l'insieme degli irrazionali di [0, 1] ha misura 1, per la (5) l'integrale di Lebesgue

$$\int_0^1 \Phi(x)\, dx$$

è uguale a 1. È facile verificare che questa funzione non ammette integrale di Riemann.

Una proposizione ausiliaria.

Sia $f(x)$ una funzione misurabile e limitata su $[a, b]$. Mostriamo che si può approssimare con precisione arbitraria mediante una combinazione lineare di funzioni caratteristiche. Siano A e B l'estremo inferiore e superiore dei suoi valori. Suddividiamo $[A, B]$ sull'asse delle ordinate mediante $y_0 = A$, y_1, ..., $y_n = B$, in intervalli di lunghezza minore di ε, dove ε è un numero positivo arbitrario. Il caso $A = B$, cioè di una funzione costante, è immediato. Se in un punto $x \in [a, b]$ vale

$$y_i \leqslant f(x) < y_{i+1} \quad (i = 0, 1, \cdots, n-1),$$

poniamo

$$\phi(x) = y_i\,,$$

e, se invece

$$f(x) = y_n = B,$$

poniamo

$$\phi(x) = y_n .$$

La costruzione di $\phi(x)$ è illustrata nella figura 3.

Per costruzione, in ogni punto di $[a, b]$ vale

$$| f(x) - \phi(x) | < \epsilon.$$

Inoltre $\phi(x)$ assume solo i valori $y_0, y_1, \ldots, y_n$; si può quindi scrivere

$$\phi(x) = y_0 \cdot \phi_0(x) + y_1 \cdot \phi_1(x) + \cdots + y_n \cdot \phi_n(x), \qquad (7)$$

dove $\phi_i(x)$ è la funzione caratteristica dell'insieme dei punti in cui $\phi(x) = y_i$, cioè $y_i \leqslant f(x) < y_{i+1}$. In ogni punto $x \in [a, b]$ una sola di queste funzioni caratteristiche vale 1 e tutte le altre valgono 0; nella somma (7) può dunque essere non nullo al più un addendo. La proposizione è dimostrata. Per $i = n$ si considera invece l'insieme dei punti per cui $f(x) = y_n$.

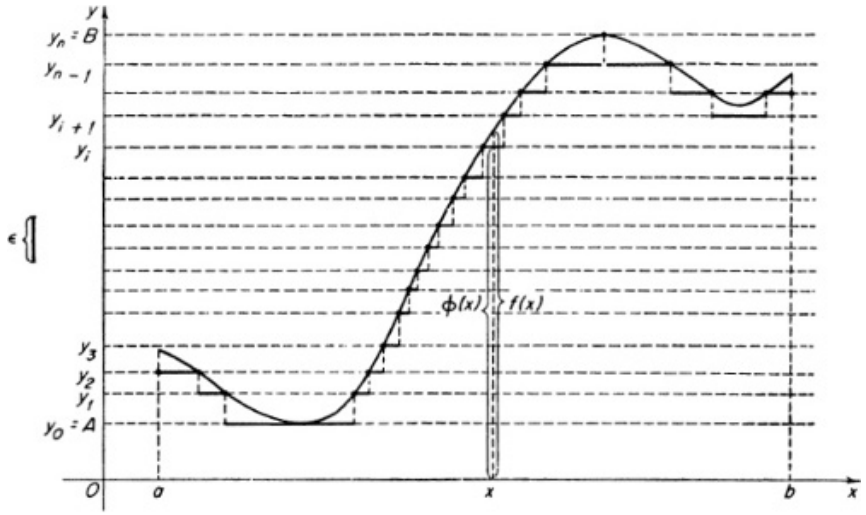


Figura 3. Approssimazione uniforme di una funzione mediante una funzione che assume un numero finito di valori.

Definizione dell'integrale di Lebesgue.

Definiamo ora l'integrale di una funzione misurabile e limitata. Poiché $\phi(x)$ differisce poco da $f(x)$, possiamo usare l'integrale della prima come approssimazione di quello della seconda. Tenendo conto che le $\phi_i(x)$ sono funzioni caratteristiche e applicando formalmente le consuete regole dell'integrazione, otteniamo

$$\begin{aligned}\int_a^b \phi(x)\,dx &= \int_a^b \{\, y_0\phi_0(x) + y_1\phi_1(x) + \cdots + y_n\phi_n(x) \,\}\,dx \\ &= y_0 \int_a^b \phi_0(x)\,dx + y_1 \int_a^b \phi_1(x)\,dx + \cdots + y_n \int_a^b \phi_n(x)\,dx \\ &= y_0\mu e_0 + y_1\mu e_1 + \cdots + y_n\mu e_n\,,\end{aligned}$$

dove μe_i è la misura dell'insieme e_i dei punti x per cui

$$y_i \le f(x) < y_{i+1} \quad \text{per } i = 0, \ldots, n-1;$$
$$f(x) = y_n = B \quad \text{per } i = n.$$

Un valore approssimato dell'integrale di Lebesgue di $f(x)$ è dunque la «somma integrale di Lebesgue»

$$S = y_0\mu e_0 + y_1\mu e_1 + \ldots + y_n\mu e_n.$$

Si definisce perciò l'integrale di Lebesgue come il limite della somma S quando

$$\max |\, y_{i+1} - y_i \,| \to 0,$$

condizione che corrisponde alla convergenza uniforme di $\phi(x)$ a $f(x)$.

Si dimostra che queste somme hanno limite per ogni funzione misurabile e limitata su $[a, b]$: ogni tale funzione è quindi integrabile secondo Lebesgue. L'integrale può essere esteso anche a certe classi di funzioni misurabili non limitate, ma qui non ce ne occuperemo.

Proprietà dell'integrale di Lebesgue.

Esso conserva le proprietà desiderabili dell'integrale ordinario: l'integrale di una somma è la somma degli integrali, un fattore costante si può portare fuori dal segno di integrale, e così via. Possiede inoltre una proprietà particolarmente importante, che l'integrale ordinario non offre con la stessa generalità: se le funzioni $f_n(x)$ sono misurabili e uniformemente limitate,

$$|\, f_n(x) | < K$$

per ogni n e ogni $x \in [a, b]$, e se $\{f_n(x)\}$ converge quasi ovunque a $f(x)$, allora

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

In altre parole, sotto queste ipotesi l'integrale di Lebesgue consente il passaggio al limite sotto il segno di integrale. Questa proprietà ne fa uno strumento molto utile e spesso indispensabile. È essenziale, in particolare, nella teoria delle serie trigonometriche, in quella degli spazi di funzioni (capitolo XIX) e in altri rami della matematica.

Facciamo un esempio. Sia $f(x)$ una funzione reale periodica di periodo 2π e sia

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

la sua serie di Fourier. Se, per esempio, $f(x)$ è continua, si dimostra che

$$\frac{1}{\pi}\int_0^{2\pi} f^2(x)\,dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty} (a_n^2 + b_n^2). \qquad (8)$$

Questa identità è detta uguaglianza di Parseval. Per quale classe di funzioni periodiche vale la (8), con entrambi i membri finiti? La risposta è: per le funzioni misurabili su $[0, 2\pi]$ il cui quadrato $f^2(x)$ è integrabile secondo Lebesgue sullo stesso intervallo, cioè le funzioni a quadrato integrabile.

Letture consigliate

F. Hausdorff, *Set theory*, 2ª ed., Chelsea, New York, 1962.

J. M. Hyslop, *Real variable*, Interscience, New York, 1960.

E. Kamke, *Theory of sets*, Dover, New York, 1950.

E. Landau, *Foundations of analysis: the arithmetic of whole, rational, irrational and complex numbers*, Chelsea, New York, 1951.

E. J. McShane e A. T. Betts, *Real analysis*, Van Nostrand, New York, 1959.

Note al capitolo XV

[1] Durante la stesura di questa sezione ci siamo avvalsi di preziose consultazioni con A. N. Kolmogorov.

[2] Lo stesso ragionamento mostra che ogni insieme numerabile di punti della retta ha misura zero.

Nota editoriale

La traduzione conserva le sei sezioni, le tre figure numerate (con le parti 2a e 2b), le cinque tavole, le formule, la bibliografia e le due note. Le figure 2a e 2b sono accostate; le loro didascalie e quelle delle altre figure sono rese esplicite in italiano. Le tavole numeriche originali sono conservate; il prospetto che enumera i razionali è ricomposto con intestazioni italiane. Si mantiene la notazione storica: somma e prodotto di insiemi indicano unione e intersezione; ϕ indica l'insieme vuoto.

Sono corretti refusi negli indici degli intervalli, nella successione usata per provare la non numerabilità, nella somma delle lunghezze di due segmenti e nel termine finale della somma integrale di Lebesgue. I simboli ε usati erroneamente al posto dell'appartenenza ∈ o del nome dell'insieme E sono corretti. Nella definizione di funzione misurabile è reinserito il segno > mancante. Nel ricoprimento dei razionali, la misura dell'unione è minore o uguale alla somma delle misure, non necessariamente minore; il rinvio alla relativa disuguaglianza è corretto. Nella dimostrazione mediante approssimazioni diadiche si usa $x_n \leq \xi$, essendo possibile l'uguaglianza.

Il divieto di appartenenza di un insieme a se stesso non costituisce, da solo, un'assiomatizzazione sufficiente a evitare tutte le antinomie. Le affermazioni generali sulla cardinalità e sugli insiemi non misurabili si intendono nell'usuale quadro della teoria degli insiemi con assioma della scelta. La parola «numerabile» indica qui un insieme infinito in corrispondenza biunivoca con i naturali; «al più numerabile» comprende anche il caso finito. Nel confronto delle cardinalità si rende esplicito che l'accoppiamento deve comprendere tutti gli elementi dell'insieme minore. Il riferimento alla dialettica è conservato come interpretazione dell'autore.

Per l'estremo superiore o inferiore si precisa che l'insieme deve essere non vuoto. Gli intervalli annidati sono chiusi e hanno lunghezze tendenti a zero, come richiede il principio di Cantor nella forma enunciata. L'equivalenza dei quattro principi di continuità presuppone la proprietà archimedea, precisazione già presente nel seguito dell'originale. Gli insiemi definiti da disuguaglianze strette per funzioni continue su [a, b] sono aperti relativamente a tale intervallo. Il caso del chiuso vuoto è incluso separatamente nella descrizione mediante intervalli contigui.

Per l'approssimazione con funzioni caratteristiche si distingue il livello finale f(x) = B dagli intervalli di livelli $y_i \leq f(x) < y_{i+1}$. In ogni punto è attiva una sola funzione caratteristica, ma il corrispondente addendo può essere zero. Sono esplicitati l'intervallo chiuso e limitato nel criterio di integrabilità di Riemann e nella formulazione del teorema di Luzin qui usata, e la possibilità di assegnare liberamente i valori su un insieme di misura nulla nei limiti quasi ovunque. Il passaggio al limite sotto il segno di integrale richiede le ipotesi enunciate di misurabilità, limitatezza uniforme e convergenza quasi ovunque: non è incondizionato. L'uguaglianza di Parseval è intesa per funzioni reali a quadrato integrabile e con membri finiti. Si conserva la convenzione dell'autore per la funzione di Dirichlet, pari a 1 sugli irrazionali e a 0 sui razionali.

CAPITOLO XVI

Algebra lineare

D. K. Faddeev

§1. Oggetto e strumenti dell'algebra lineare

Funzioni lineari e matrici.

Fra le funzioni di una variabile, la più semplice è la cosiddetta funzione lineare $l(x) = ax + b$. Il suo grafico è naturalmente la più semplice delle curve: una retta.

Eppure la funzione lineare è fra le più importanti. Infatti, un piccolo tratto di una curva «liscia» assomiglia a un segmento di retta, tanto più quanto minore è la sua curvatura. Nel linguaggio delle funzioni, una funzione «liscia», cioè continuamente derivabile, per piccole variazioni della variabile indipendente è vicina a una funzione lineare. Quest'ultima è caratterizzata dalla proporzionalità fra il suo incremento e quello della variabile indipendente. Infatti, $\Delta l = l(x_0 + \Delta x) - l(x_0) = a(x_0 + \Delta x) + b - (ax_0 + b) = a\Delta x$. Viceversa, se $\Delta l = a\Delta x$, allora $l(x) - l(x_0) = a(x - x_0)$, quindi $l(x) = ax + l(x_0) - ax_0 = ax + b$, con $b = l(x_0) - ax_0$. Dal calcolo differenziale sappiamo che nell'incremento di una funzione derivabile si distingue una parte principale, il differenziale, proporzionale all'incremento della variabile indipendente; la differenza fra incremento e differenziale è un infinitesimo di ordine superiore. Per una variazione infinitesima della variabile, una funzione derivabile è dunque approssimata da una funzione lineare a meno di un infinitesimo di ordine superiore.

Per le funzioni di più variabili la situazione è analoga.

Una funzione lineare di più variabili ha la forma $a_1x_1 + a_2x_2 + \ldots + a_nx_n + b$. Se $b = 0$, si dice omogenea. È caratterizzata dalle due proprietà seguenti.

1. Se varia una sola variabile indipendente e le altre restano costanti, l'incremento della funzione è proporzionale a quello della variabile.

2. Se variano tutte le variabili, l'incremento della funzione è la somma algebrica degli incrementi ottenuti facendo variare ciascuna variabile separatamente.

La funzione lineare di più variabili svolge lo stesso ruolo che quella di una variabile svolge fra le funzioni di una variabile. Una funzione «liscia», cioè con derivate parziali continue rispetto a tutte le variabili, per piccole variazioni di queste è vicina a una funzione lineare. L'incremento di $w = f(x_1, \ldots, x_n)$ coincide, a meno di infinitesimi di ordine superiore, con il differenziale totale $(\partial f/\partial x_1)\, dx_1 + \ldots + (\partial f/\partial x_n)\, dx_n$, funzione lineare omogenea degli incrementi $dx_1, \ldots, dx_n$. La funzione stessa, somma del valore iniziale e dell'incremento, si

esprime quindi localmente come una funzione lineare non omogenea, a meno di infinitesimi di ordine superiore.

Lo studio della dipendenza di una grandezza da più fattori conduce a funzioni di più variabili. Un problema si dice lineare se tale dipendenza è lineare. Le proprietà appena viste permettono di caratterizzarlo come segue.

1. Proporzionalità: l'effetto di ciascun fattore, considerato separatamente, è proporzionale al suo valore.

2. Indipendenza: l'effetto complessivo è la somma degli effetti dei singoli fattori.

La possibilità di sostituire in prima approssimazione una funzione liscia con una lineare, per piccole variazioni delle variabili, riflette un principio generale: un problema di variazione di una grandezza sotto l'azione di più fattori, per piccole azioni e in prima approssimazione, può essere trattato come lineare, cioè dotato di proporzionalità e indipendenza. Spesso questa approssimazione basta agli scopi pratici, come nella teoria classica dell'elasticità e in quella delle piccole oscillazioni.

Le grandezze fisiche sono spesso descritte da più numeri: una forza dalle tre proiezioni sugli assi, lo stato di tensione in un punto di un corpo elastico dalle sei componenti del tensore degli sforzi, ecc. Occorre allora considerare simultaneamente più funzioni di più variabili e, in prima approssimazione, più funzioni lineari.

Una funzione lineare di una variabile è tanto semplice da non richiedere una teoria particolare. Con più variabili compaiono invece caratteristiche nuove; il problema si complica ulteriormente passando da una funzione delle variabili $x_1, \ldots, x_n$ a più funzioni $y_1, \ldots, y_m$ delle stesse variabili. In prima approssimazione si presenta un sistema di funzioni lineari:

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n + b_1 , \\ y_2 &= a_{21}x_1 + \cdots + a_{2n}x_n + b_2 , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n + b_m . \end{aligned}$$

Un sistema di funzioni lineari è già un oggetto matematico abbastanza complesso, il cui studio offre risultati interessanti e non banali.

Lo studio delle funzioni lineari e dei loro sistemi costituisce il punto di partenza del ramo dell'algebra chiamato algebra lineare.

Storicamente, il suo primo problema è la risoluzione di un sistema di equazioni lineari:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 , \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 , \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m . \end{aligned}$$

Il caso più semplice si studia nell'algebra elementare. La ricerca di metodi numerici semplici ed efficienti per sistemi con *n* grande continua a richiamare l'attenzione dei ricercatori, perché la loro soluzione entra come componente essenziale in molti calcoli e indagini.

Le funzioni lineari omogenee si chiamano anche forme lineari. Un sistema di forme lineari

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n\,, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n \end{aligned}$$

è completamente descritto dai suoi coefficienti: le proprietà dipendono soltanto dai loro valori numerici, mentre i nomi delle variabili sono irrilevanti.

Per esempio, i sistemi di forme

$$\begin{gathered} 3x_1 + x_2 - x_3, \quad 3t_1 + t_2 - t_3, \\ 2x_1 + x_2 + 3x_3, \text{ e } 2t_1 + t_2 + 3t_3, \\ x_1 - x_2 - x_3, \quad t_1 - t_2 - t_3. \end{gathered}$$

hanno evidentemente le stesse proprietà e non devono essere considerati sostanzialmente diversi.

I coefficienti di un sistema di forme lineari si dispongono naturalmente in una tabella rettangolare:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

Queste tabelle si chiamano matrici e i numeri a_{ij} sono i loro elementi. La necessità delle matrici nasce dunque dall'oggetto stesso dell'algebra lineare.

Casi particolari importanti sono le matrici con una sola colonna, dette colonne; quelle con una sola riga, dette righe; e le matrici quadrate, con uguale numero di righe e colonne. Il numero delle righe di una matrice quadrata si chiama ordine. La matrice (a), formata da un solo numero, si identifica con quel numero.

Le operazioni più semplici sui sistemi di forme suggeriscono naturalmente operazioni sulle matrici.

Siano dati i due sistemi di forme lineari

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n\,, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n\,, \end{aligned}$$

e

$$\begin{aligned} z_1 &= b_{11}x_1 + \cdots + b_{1n}x_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_m &= b_{m1}x_1 + \cdots + b_{mn}x_n . \end{aligned}$$

Sommiamo le forme corrispondenti:

$$\begin{aligned} y_1 + z_1 &= (a_{11} + b_{11})\, x_1 + \cdots + (a_{1n} + b_{1n})\, x_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m + z_m &= (a_{m1} + b_{m1})\, x_1 + \cdots + (a_{mn} + b_{mn})\, x_n . \end{aligned}$$

È naturale chiamare la matrice del sistema ottenuto,

$$\begin{bmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{bmatrix}$$

somma delle matrici

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

Analogamente, il prodotto della matrice

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

per il numero c si definisce come la matrice dei coefficienti delle forme $cy_1, \ldots, cy_m$, dove $y_1, \ldots, y_m$ hanno per matrice dei coefficienti

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} .$$

Dalla definizione risulta che

$$c \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & \cdots & ca_{1n} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ ca_{m1} & \cdots & ca_{mn} \end{bmatrix} .$$

Il prodotto di due matrici si definisce infine nel modo seguente. Siano

$$\begin{aligned} z_1 &= a_{11}y_1 + \cdots + a_{1m}y_m , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_k &= a_{k1}y_1 + \cdots + a_{km}y_m \end{aligned} \qquad (1)$$

e

$$\begin{aligned} y_1 &= b_{11}x_1 + \cdots + b_{1n}x_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= b_{m1}x_1 + \cdots + b_{mn}x_n . \end{aligned}$$

Sostituendo nella (1) le espressioni delle y_1, ..., y_m in funzione delle x_1, ..., x_n, si esprimono anche z_1, ..., z_k mediante forme lineari delle stesse variabili:

$$\begin{aligned} z_1 &= c_{11}x_1 + \cdots + c_{1n}x_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_k &= c_{k1}x_1 + \cdots + c_{kn}x_n . \end{aligned}$$

La matrice dei coefficienti

$$\begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{k1} & \cdots & c_{kn} \end{bmatrix}$$

si chiama prodotto delle matrici

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$

e si indica con

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{km} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix} .$$

È facile esprimere gli elementi del prodotto in funzione degli elementi dei fattori. L'elemento c_{ij} è il coefficiente di x_j nell'espressione di z_i in funzione delle x_1, ..., x_n.

Ma $z_i = a_{i1}y_1 + \ldots + a_{im}y_m$ e

$$\begin{aligned} y_1 &= \cdots + b_{1j}x_j + \cdots , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m &= \cdots + b_{mj}x_j + \cdots . \end{aligned}$$

Pertanto

$$z_i = \cdots + (a_{i1}b_{1j} + \cdots + a_{im}b_{mj})\, x_j + \cdots ,$$

da cui

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{im}b_{mj} .$$

L'elemento della riga *i* e della colonna *j* del prodotto è dunque la somma dei prodotti degli elementi della riga *i* del primo fattore per i corrispondenti elementi della colonna *j* del secondo. Per esempio:

$$\begin{pmatrix} 2 & 1 & 3 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2\cdot 3 + 1\cdot 1 + 3\cdot 2 & 2\cdot 1 + 1\cdot 2 + 3\cdot 4 \\ 3\cdot 3 + 1\cdot 1 + 1\cdot 2 & 3\cdot 1 + 1\cdot 2 + 1\cdot 4 \end{pmatrix} = \begin{pmatrix} 13 & 16 \\ 12 & 9 \end{pmatrix}.$$

Benché una matrice sia un oggetto «composto» da molti elementi, è utile indicarla con una sola lettera e conservare le consuete notazioni per somma e prodotto. Useremo lettere latine maiuscole. Questa notazione compatta rende la teoria più semplice e chiara: formule brevi, simili a quelle dell'algebra ordinaria, esprimono relazioni complesse fra tutti gli elementi delle matrici. Così il sistema di forme

$$\begin{aligned} &a_{11}x_1 + \cdots + a_{1n}x_n, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{m1}x_1 + \cdots + a_{mn}x_n \end{aligned}$$

si scrive AX, dove A è la matrice dei coefficienti e X la colonna delle variabili x_1, ..., x_n. Il sistema di equazioni

$$\begin{aligned} &a_{11}x_1 + \cdots + a_{1n}x_n = b_1, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{aligned}$$

si scrive

$$AX = B,$$

dove A è la matrice dei coefficienti, X la colonna delle incognite e B quella dei termini noti.

Somma e prodotto di matrici non sono sempre definiti. Si possono sommare matrici con lo stesso numero di righe e colonne; il risultato ha le stesse dimensioni. Il prodotto è definito se il numero di colonne del primo fattore coincide con il numero di righe del secondo. La matrice risultante ha tante righe quante il primo fattore e tante colonne quante il secondo.

Le operazioni sulle matrici quadrate soddisfano molte delle leggi valide per i numeri, ma alcune di queste leggi non valgono più.

Elenchiamo le proprietà fondamentali delle operazioni sulle matrici:

1. $A + B = B + A$ (proprietà commutativa della somma).
2. $(A + B) + C = A + (B + C)$ (proprietà associativa della somma).
3. $c(A + B) = cA + cB$.
3′. $(c_1 + c_2)A = c_1A + c_2A$.
Le ultime due sono le proprietà distributive del prodotto per un numero; c, c_1 e c_2 sono numeri, non matrici.

4. $(c_1c_2)A = c_1(c_2A)$ (proprietà associativa del prodotto per un numero).

5. Esiste una matrice nulla

$$O = \begin{pmatrix} 0 \cdots 0 \\ \cdots\cdots \\ 0 \cdots 0 \end{pmatrix}$$

tale che $A + O = A$ per ogni A delle stesse dimensioni.

6. $cO = 0A = O$; viceversa, se $cA = O$, allora $c = 0$ oppure $A = O$, dove c è un numero.

7. Per ogni matrice A esiste l'opposta $-A$, tale che $A + (-A) = O$.

8. $(A + B)C = AC + BC$.
8′. $C(A + B) = CA + CB$.
Sono le proprietà distributive del prodotto di matrici rispetto alla somma.
9. $(AB)C = A(BC)$ (proprietà associativa del prodotto).
10. $(cA)B = A(cB) = c(AB)$.

Queste proprietà valgono anche per matrici rettangolari, purché tutte le operazioni indicate siano definite. Per matrici quadrate dello stesso ordine questa condizione è automaticamente soddisfatta.

Le proprietà elencate sono analoghe a quelle delle operazioni sui numeri.

Segnaliamo ora due particolarità. Anzitutto, il prodotto di matrici, anche quadrate, non è in generale commutativo: AB può essere diverso da BA. Per esempio:

$$\begin{pmatrix} 3 & -2 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} -3 & 2 \\ 11 & 6 \end{pmatrix};$$

$$\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 6 \\ 7 & 2 \end{pmatrix}.$$

Inoltre, il prodotto di due numeri è zero se e solo se almeno un fattore è zero, proprietà fondamentale nella teoria delle equazioni algebriche. Per le matrici questo non vale: il prodotto può essere nullo senza che lo sia alcuno dei fattori. Per esempio:

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ricordiamo un'altra proprietà del prodotto. La matrice $\bar{A}$ si chiama trasposta di A quando ogni sua riga contiene, nello stesso ordine, gli elementi della corrispondente colonna di A. Per esempio, la matrice

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

ha per trasposta

$$\bar{A} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

Prodotto e trasposizione sono legati dalla formula

$$\overline{AB} = \bar{B}\bar{A} ,$$

che si verifica facilmente usando la regola del prodotto.

La teoria delle matrici è uno strumento indispensabile dell'algebra lineare per formulare e risolvere i suoi problemi.

Analogie geometriche nell'algebra lineare.

Oltre alle origini già descritte, anche le esigenze dell'analisi e della geometria, soprattutto analitica, hanno promosso l'algebra lineare, arricchendola di idee e analogie. La geometria analitica del piano e ancor più dello spazio, nello studio di rette e piani, usa infatti l'algebra lineare nelle sue forme più semplici. Una retta nel piano è descritta da un'equazione lineare fra le due coordinate dei suoi punti; un piano nello spazio da un'equazione lineare in tre coordinate; una retta nello spazio da due equazioni lineari.

Il concetto di vettore rende particolarmente semplice e chiara la geometria analitica e quindi la teoria dei sistemi lineari più semplici. Un concetto generalizzato di vettore conferisce la stessa chiarezza all'algebra lineare e alla teoria generale dei sistemi. La generalizzazione nasce così: un vettore dello spazio è determinato da tre numeri, le sue proiezioni sugli assi; viceversa, ogni terna di reali si rappresenta geometricamente con un vettore.

Per i vettori sono definite la somma, mediante la regola del parallelogramma, e la moltiplicazione per un numero. Queste operazioni corrispondono a quelle sulle forze, velocità, accelerazioni e altre grandezze fisiche vettoriali.

Se i vettori sono dati mediante le coordinate, cioè le proiezioni sugli assi, somma e moltiplicazione per un numero corrispondono alle analoghe operazioni sulle righe o colonne delle coordinate.

È quindi naturale interpretare una riga o colonna di tre elementi come un vettore dello spazio tridimensionale. Le operazioni sulle righe o colonne si interpretano come operazioni sui vettori; formalmente le due algebre coincidono. Ciò suggerisce di introdurre una terminologia geometrica nell'algebra lineare.

Una colonna o riga di *n* numeri

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

viene considerata un «vettore», cioè un elemento di uno «spazio vettoriale a *n* dimensioni». La somma dei vettori

$$\boxed{\begin{matrix} x_1 \\ x_2 \\ x_{\cdots} \\ x_n \end{matrix}} \qquad \text{e} \qquad \boxed{\begin{matrix} y_1 \\ y_2 \\ y_{\cdots} \\ y_n \end{matrix}}$$

è per definizione il vettore

$$\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \cdots\cdots \\ x_n + y_n \end{pmatrix};$$

e il prodotto del vettore

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

per il numero *c* è

$$\begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}.$$

L'insieme di tutti questi vettori, o colonne, costituisce per definizione lo spazio vettoriale aritmetico a *n* dimensioni.

Accanto a esso possiamo introdurre uno spazio di punti a *n* dimensioni, associando a ogni colonna di *n* reali un punto. A ogni coppia di punti *A*, *B* associamo allora il vettore $\overrightarrow{AB}$, diretto da *A* a *B*, le cui coordinate sono le differenze fra le coordinate corrispondenti di *B* e di *A*. Due

vettori sono uguali se hanno coordinate uguali, proprio come nello spazio tridimensionale si considerano uguali vettori ottenibili l'uno dall'altro mediante una traslazione.

Fra i vettori e i punti degli spazi così definiti esiste una corrispondenza biunivoca naturale.

Assumiamo il punto

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

come origine e associamo a ogni punto il vettore che va dall'origine a esso. Viceversa, a ogni vettore applicato nell'origine associamo il suo estremo. Lo spazio di punti offre nuove analogie che aiutano a «vedere» meglio lo spazio a n dimensioni.

Nelle generalizzazioni successive (§2), però, definire rigorosamente lo spazio di punti richiede qualche complicazione ulteriore; non useremo perciò questo concetto. Il lettore che voglia ricorrere a tali analogie può rappresentare gli elementi dello spazio vettoriale come vettori uscenti dall'origine.

La terminologia geometrica permette di usare nell'algebra lineare analogie fondate sull'intuizione maturata studiando la geometria tridimensionale.

Occorre naturalmente prudenza: ogni argomento geometrico intuitivo deve poter essere verificato logicamente mediante definizioni precise e dimostrazioni rigorose.

Gli elementi dello spazio vettoriale a n dimensioni possiedono somma e moltiplicazione per un numero, con proprietà simili a quelle numeriche. La somma è commutativa e associativa, valgono le leggi distributive per la moltiplicazione scalare, la sottrazione è univocamente definita e il prodotto di un vettore per un numero è nullo se e solo se è nullo il vettore oppure il numero.

Queste proprietà non appartengono soltanto a righe e colonne. Valgono per le matrici delle stesse dimensioni e per grandezze fisiche come forze, velocità e accelerazioni. Valgono anche per oggetti del tutto diversi: tutti i polinomi in una variabile, tutte le funzioni continue su $[a, b]$, tutte le soluzioni di un'equazione differenziale lineare omogenea, e così via.

Ciò motiva un'ulteriore generalizzazione: gli spazi lineari astratti. I loro elementi possono essere oggetti matematici o fisici arbitrari sui quali siano definite naturalmente somma e moltiplicazione per un numero. Questa impostazione generale non complica la teoria: ogni spazio lineare di dimensione finita n, concetto che preciseremo nella prossima sezione, ha la stessa struttura e le stesse proprietà dello spazio aritmetico corrispondente. La generalizzazione amplia però enormemente le applicazioni, consentendo di usare l'algebra lineare in molti problemi delle scienze teoriche.

§2. Spazi lineari

Definizione di spazio lineare.

Passiamo a una definizione rigorosa.

Uno spazio lineare è un insieme di oggetti di natura arbitraria per i quali siano definite una somma e una moltiplicazione per un numero, soggette ai postulati seguenti:

1. $(X + Y) + Z = X + (Y + Z)$.

2. Esiste un elemento nullo 0 tale che $X + 0 = X$ per ogni X.

3. Per ogni X esiste un opposto $-X$ tale che $X + (-X) = 0$.

4. $X + Y = Y + X$.

5. $1 \cdot X = X$.

6. $c_1(c_2X) = (c_1c_2)X$.

7. $(c_1 + c_2)X = c_1X + c_2X$.

8. $c(X + Y) = cX + cY$.

Qui X, Y, Z sono elementi dello spazio; 1, c_1, c_2, c sono numeri.

Questi postulati, detti anche assiomi dello spazio lineare, formalizzano le proprietà che associamo alla somma e alla moltiplicazione per un numero anche nel loro significato più generale. Operazioni con significati fisici diversi vengono trattate come somma e moltiplicazione scalare proprio quando soddisfano gli assiomi 1-8.

Ne seguono alcune conseguenze:

a. L'elemento nullo è unico: un solo elemento soddisfa l'assioma 2.

b. L'opposto di un elemento X è unico.

c. La sottrazione ha significato: data una somma e uno degli addendi, l'altro è determinato univocamente. Se $X + Z = Y$, allora $Z = Y + (-X)$.

d. $0 \cdot X = c \cdot 0 = 0$.

e. Se $cX = 0$, allora $c = 0$ oppure $X = 0$.

f. $-X = (-1)X$.

Omettiamo le dimostrazioni, molto semplici. D'ora in poi chiameremo vettori gli elementi dello spazio lineare.

Dipendenza e indipendenza lineare.

Introduciamo questi concetti fondamentali.

Il vettore $c_1X_1 + \ldots + c_mX_m$, con coefficienti numerici arbitrari, si chiama combinazione lineare di $X_1, \ldots, X_m$. Se almeno uno dei vettori è combinazione lineare degli altri, essi si dicono linearmente dipendenti; altrimenti si dicono linearmente indipendenti.

È facile vedere che l'indipendenza equivale alla seguente condizione: $c_1X_1 + \ldots + c_mX_m = 0$ soltanto quando tutti i coefficienti $c_1, \ldots, c_m$ sono nulli.

Nello spazio tridimensionale ordinario questi concetti hanno un significato geometrico semplice.

Due vettori X_1 e X_2 sono dipendenti se uno è un multiplo numerico dell'altro. Se non sono nulli, appartengono quindi alla stessa retta e hanno direzioni concordi o opposte.

Viceversa, vettori sulla stessa retta sono dipendenti. Due vettori indipendenti non possono dunque essere collocati sulla stessa retta: le loro direzioni sono diverse.

Consideriamo tre vettori dipendenti e supponiamo, per fissare le idee, che X_3 sia combinazione di X_1 e X_2. Esso appartiene allora a un piano che contiene gli altri due. Viceversa, tre vettori complanari sono dipendenti: se X_1 e X_2 non sono allineati, X_3 si decompone lungo di essi; se sono allineati, i primi due sono già dipendenti.

La dipendenza di tre vettori equivale dunque alla loro complanarità; essi sono indipendenti se e solo se non appartengono tutti allo stesso piano.

Quattro vettori dello spazio tridimensionale sono sempre dipendenti. Se i primi tre lo sono già, lo sono anche insieme al quarto. Se invece sono indipendenti, non sono complanari e ogni quarto vettore si esprime come loro combinazione lineare.

Queste osservazioni si generalizzano nel modo seguente.

Nello spazio tridimensionale i vettori $X_1, \ldots, X_k$, con $k \leq 3$, sono dipendenti se e solo se appartengono a un sottospazio di dimensione minore di k: il sottospazio nullo, una retta o un piano, secondo il caso.

Dopo avere definito rigorosamente sottospazio e dimensione, vedremo che questa caratterizzazione vale in generale. Il significato geometrico della dipendenza rimane dunque quello osservato nello spazio tridimensionale.

Il teorema seguente è fondamentale: se $X_1, \ldots, X_m$ sono combinazioni lineari di $Y_1, \ldots, Y_k$ e $m > k$, allora i vettori $X_1, \ldots, X_m$ sono dipendenti. Lo chiameremo teorema sulla dipendenza delle combinazioni lineari.

Per $k = 1$ è evidente; per $k > 1$ si dimostra facilmente per induzione rispetto a k.

Base e dimensione.

Nello spazio tridimensionale tre vettori non complanari, cioè indipendenti, formano una base: ogni vettore si decompone lungo di essi, ossia si esprime come loro combinazione lineare.

Gli spazi lineari generali si dividono in due tipi.

Uno spazio può contenere famiglie di vettori indipendenti arbitrariamente numerose. Si dice allora di dimensione infinita; il suo studio conduce a problemi trattati dall'analisi funzionale, cui è dedicato il capitolo XIX.

Uno spazio è di dimensione finita se il numero dei vettori indipendenti ha un limite finito: esistono n vettori indipendenti, ma ogni famiglia con più di n vettori è dipendente. Il numero n si chiama dimensione dello spazio.

Lo spazio dei vettori della geometria ordinaria ha quindi dimensione tre anche secondo questa definizione: contiene terne indipendenti, mentre ogni gruppo di quattro vettori è dipendente.

Lo spazio delle colonne di n elementi ha dimensione n. Vi sono infatti n vettori indipendenti, per esempio

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \qquad (2)$$

e ogni vettore

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

è la combinazione lineare $x_1 e_1 + \ldots + x_n e_n$. Il teorema precedente implica allora che più di n vettori sono sempre dipendenti.

I polinomi in una variabile formano uno spazio lineare: somma e moltiplicazione scalare sono definite naturalmente e soddisfano gli assiomi 1-8. Lo spazio è infinito-dimensionale, perché $1, x, \ldots, x^N$ sono indipendenti per ogni N. I polinomi di grado non superiore a un dato N, compreso il polinomio nullo, formano invece uno spazio di dimensione $N + 1$. I vettori $1, x, \ldots, x^N$ sono infatti indipendenti e ogni polinomio di quel grado è loro combinazione; pertanto più di $N + 1$ polinomi di grado $\leq N$ sono dipendenti.

Una base di uno spazio è una famiglia di vettori indipendenti che permette di esprimere ogni vettore dello spazio come combinazione lineare. Nello spazio delle colonne, per esempio, i vettori (2) formano una base. Per i polinomi di grado $\leq N$ una base è $1, x, \ldots, x^N$. Nello spazio tridimensionale ogni terna indipendente è una base.

In uno spazio di dimensione n, ogni famiglia di n vettori indipendenti è una base. Siano infatti $e_1, \ldots, e_n$ indipendenti e X arbitrario. La famiglia $X, e_1, \ldots, e_n$ è dipendente perché contiene più di n elementi. Esistono dunque coefficienti non tutti nulli tali che $cX + c_1e_1 + \ldots + c_ne_n = 0$. Necessariamente $c \neq 0$, altrimenti i vettori e_i sarebbero dipendenti. Quindi $X = -(c_1/c)e_1 - \ldots - (c_n/c)e_n$, come volevamo.

Ogni base di uno spazio di dimensione n contiene esattamente n vettori. Essendo indipendenti, i suoi k vettori soddisfano $k \leq n$. D'altra parte ogni vettore dello spazio è loro combinazione: per il teorema sulla dipendenza delle combinazioni lineari, più di k vettori sono dipendenti. La dimensione n non può quindi superare k, e risulta $k = n$.

Introduciamo le coordinate rispetto a una base $e_1, \ldots, e_n$. Ogni vettore X è una combinazione dei vettori della base e la rappresentazione è unica. Supponiamo infatti di averne due:

$$X = x_1e_1 + x_2e_2 + \cdots + x_ne_n ,$$

$$X = x'_1e_1 + x'_2e_2 + \cdots + x'_ne_n .$$

Allora

$$(x_1 - x'_1)\, e_1 + (x_2 - x'_2)\, e_2 + \cdots + (x_n - x'_n)\, e_n = 0$$

e l'indipendenza dei vettori della base implica $x_1 = x'_1, \ldots, x_n = x'_n$.

I coefficienti $x_1, \ldots, x_n$ della decomposizione si chiamano coordinate di X nella base scelta. A ogni vettore associamo così la riga o colonna delle sue coordinate; viceversa, ogni riga o colonna di n numeri individua un vettore.

Somma e moltiplicazione scalare dei vettori corrispondono alle stesse operazioni sulle loro coordinate.

Per queste operazioni, ogni spazio di dimensione n, qualunque sia la natura dei suoi elementi, funzioni, matrici o grandezze fisiche, ha dunque la stessa struttura dello spazio delle righe o colonne. L'impostazione assiomatica non complica la teoria, ma ne amplia considerevolmente le applicazioni.

L'identità di struttura di due insiemi di oggetti rispetto a operazioni o relazioni assegnate si chiama isomorfismo; una definizione precisa sarà data nel capitolo XX. Possiamo dunque dire che tutti gli spazi lineari di dimensione n sullo stesso campo sono isomorfi tra loro e al modello dello spazio delle righe.

Sottospazi.

Un sottoinsieme non vuoto di uno spazio lineare R_n, chiuso rispetto a tutte le combinazioni lineari dei suoi elementi, si chiama sottospazio. È esso stesso uno spazio lineare e possiede basi e dimensione. La sua dimensione non supera quella dello spazio ambiente e coincide con essa soltanto se il sottospazio è l'intero spazio.

Nello spazio tridimensionale sono esempi di sottospazi le rette e i piani passanti per l'origine, o, equivalentemente, gli insiemi dei vettori contenuti in tali rette e piani.

Spesso si studia il sottospazio generato da alcuni vettori. Dati X_1, ..., X_m, dipendenti o indipendenti, tutte le combinazioni $c_1X_1 + \ldots + c_mX_m$ formano un sottospazio, detto generato da quei vettori.

La sua dimensione è il rango della famiglia X_1, ..., X_m. È facile vedere che il rango coincide con il massimo numero di vettori indipendenti estraibili dalla famiglia.

Anche l'insieme formato dal solo vettore nullo è un sottospazio; gli si attribuisce dimensione zero.

Da due sottospazi di R_n si costruiscono naturalmente altri due sottospazi: la somma vettoriale e l'intersezione.

La somma vettoriale di P e Q è l'insieme di tutte le somme di un vettore di P e uno di Q. È anche il sottospazio generato dall'unione di due basi dei sottospazi.

L'intersezione consiste dei vettori comuni. Per esempio, nello spazio tridimensionale la somma di due piani vettoriali distinti è l'intero spazio e la loro intersezione è una retta.

Se p e q sono le dimensioni dei due sottospazi, t quella della somma e s quella dell'intersezione, vale la relazione

$$p + q = t + s.$$

della quale omettiamo la dimostrazione.

Ne seguono conseguenze interessanti. In uno spazio a quattro dimensioni, due piani distinti si intersecano in generale soltanto nel vettore nullo, cioè in un sottospazio di dimensione zero. Si intersecano lungo una retta soltanto se la loro somma ha dimensione tre, ossia se entrambi sono contenuti in un sottospazio tridimensionale. Infatti $t + s = 2 + 2 = 4$, quindi $s = 1$ soltanto per $t = 3$.

Spazi lineari complessi.

Finora non abbiamo precisato quali numeri siano ammessi nella moltiplicazione scalare. Partendo dai vettori geometrici ordinari, avevamo in mente numeri reali arbitrari; si ottengono così gli spazi lineari reali. In molti problemi è però utile considerare spazi complessi: insiemi dotati di somma e moltiplicazione per un numero complesso arbitrario che soddisfano gli stessi assiomi 1-8. Ne è un esempio lo spazio delle righe con elementi complessi.

Formalmente, la teoria degli spazi complessi non differisce in modo essenziale da quella degli spazi reali.

Già uno spazio complesso di dimensione due non ha però una rappresentazione nella geometria ordinaria tridimensionale. Uno spazio complesso di dimensione n può essere considerato anche reale, limitando la moltiplicazione scalare ai numeri reali; la sua dimensione reale è allora $2n$. Se $e_1, \ldots, e_n$ è una base complessa, una base reale è data, per esempio, da

$$e_1, ie_1, e_2, ie_2, \ldots, e_n, ie_n, \quad \text{dove } i^2 = -1.$$

Uno spazio complesso bidimensionale può quindi essere interpretato come uno spazio reale quadrimensionale.

La teoria conserva formalmente la stessa forma se gli scalari appartengono a un campo numerico diverso dai reali o dai complessi, purché siano ammesse somma, sottrazione, moltiplicazione e divisione per elementi non nulli e i risultati restino nel campo. Questo concetto sarà approfondito nel capitolo XX. Un esempio è il campo dei razionali.

La teoria degli spazi lineari su un campo arbitrario trova applicazioni importanti in parti dell'algebra vicine alla teoria dei numeri.

Lo spazio euclideo a *n* dimensioni.

Non abbiamo ancora generalizzato alcuni concetti geometrici essenziali, in particolare lunghezza e angolo. In geometria analitica, intersezioni, parallelismo e molti altri problemi non ne fanno uso. Le proprietà che non richiedono lunghezze e angoli rimangono invariate sotto trasformazioni affini arbitrarie (capitolo III, §11). Per questo l'originale chiama «affini» gli spazi lineari considerati senza una struttura metrica.

Molti problemi richiedono però di estendere lunghezze e angoli agli spazi a n dimensioni. Si procede per analogia con i vettori del piano e dello spazio.

Nello spazio reale delle righe, definiamo la lunghezza di $X = (x_1, \ldots, x_n)$ come

$$|X| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

È naturale: per $n = 2$ e $n = 3$ questa è precisamente la formula cartesiana della lunghezza.

Nel piano e nello spazio, l'angolo fra X e Y è l'angolo in A di un triangolo con lati $AB = |X|$, $AC = |Y|$ e $BC = |X - Y|$.

Sarebbe naturale assumerlo come definizione anche in dimensione n, come se potessimo disegnare i due vettori e collocarli in un piano conservando lunghezze e angolo. Manca però ancora una giustificazione: bisogna provare che esiste un triangolo con quei tre lati, eventualmente degenere.

Trascuriamo per il momento questo punto e ricaviamo la formula dell'angolo. Per il teorema del coseno,

$$BC^2 = AB^2 + AC^2 - 2AB \cdot AC \cos\phi;$$

dunque

$$\begin{aligned}\cos\phi &= \frac{|X|^2 + |Y|^2 - |X - Y|^2}{2|X| \cdot |Y|} \\ &= \frac{x_1^2 + \cdots + x_n^2 + y_1^2 + \cdots + y_n^2 - (x_1 - y_1)^2 - \cdots - (x_n - y_n)^2}{2|X| \cdot |Y|} \\ &= \frac{x_1 y_1 + \cdots + x_n y_n}{|X| \cdot |Y|}.\end{aligned}$$

Conservando il nome di prodotto scalare per il prodotto delle lunghezze per il coseno dell'angolo, otteniamo

$$X \cdot Y = x_1 y_1 + \cdots + x_n y_n,$$

che per $n = 2$ e 3 coincide con le formule note per i vettori ordinari.

A rigore, conviene definire prima il prodotto scalare mediante $x_1y_1 + \ldots + x_ny_n$, poiché la definizione tramite l'angolo non è ancora giustificata, e definire poi l'angolo fra vettori non nulli con

$$\cos\phi = \frac{X \cdot Y}{|X| \cdot |Y|}. \qquad (3)$$

Seguiremo proprio questa strada.

Per giustificare la definizione bisogna mostrare che il valore assoluto del secondo membro della (3) non supera 1, cioè $(X \cdot Y)^2 \leqslant |X|^2 \cdot |Y|^2$.

Scritta per esteso, questa disuguaglianza diventa

$$(x_1y_1 + \ldots + x_ny_n)^2 \leq (x_1^2 + \ldots + x_n^2)(y_1^2 + \ldots + y_n^2).$$

È la disuguaglianza di Cauchy-Bunjakovskij, nota anche come disuguaglianza di Cauchy-Schwarz. Si può provarla direttamente con un calcolo piuttosto lungo; useremo invece il ragionamento seguente.

Il prodotto scalare possiede anzitutto queste proprietà:

1′. $X \cdot X = |X|^2 > 0$ se $X \neq 0$.

2′. $X \cdot Y = Y \cdot X$.

3′. $(cX) \cdot Y = c(X \cdot Y)$.

4′. $(X_1 + X_2) \cdot Y = X_1 \cdot Y + X_2 \cdot Y$.

Esse seguono immediatamente dall'espressione in coordinate.

Consideriamo $Y + tX$, con t reale arbitrario. Vale $|Y + tX|^2 \geqslant 0$, poiché il quadrato di una lunghezza non è negativo. Ma $|Y + tX|^2 = |Y|^2 + 2t(X \cdot Y) + t^2|X|^2$. Se $X \neq 0$, questo trinomio ha coefficiente principale positivo; è non negativo per ogni t se e solo se il discriminante è non positivo. Il discriminante è $4(X \cdot Y)^2 - 4|X|^2|Y|^2$, quindi $(X \cdot Y)^2 - |X|^2|Y|^2 \leqslant 0$, equivalente alla disuguaglianza cercata. Se $X = 0$, il risultato è immediato.

Ne segue che

$$\frac{|X \cdot Y|}{|X||Y|} \leqslant 1$$

e la definizione dell'angolo mediante la (3) è giustificata.

Si ricavano inoltre facilmente le disuguaglianze

$$|X| - |Y| \leqslant |X \pm Y| \leqslant |X| + |Y|,$$

che garantiscono l'esistenza di un triangolo, eventualmente degenere, con lati $|X|$, $|Y|$ e $|X - Y|$. È così giustificata anche la precedente interpretazione geometrica.

Definizione assiomatica di spazio euclideo.

Abbiamo introdotto lunghezza, angolo e prodotto scalare nello spazio delle righe. In uno spazio lineare reale astratto si procede assiomaticamente, ponendo al primo posto il prodotto scalare.

Un prodotto scalare associa a ogni coppia X, Y un numero reale $X \cdot Y$, soddisfacendo i seguenti assiomi:

1′. $X \cdot X > 0$ se $X \neq 0$; $0 \cdot 0 = 0$.

2′. $X \cdot Y = Y \cdot X$.

3′. $(cX) \cdot Y = c(X \cdot Y)$.

4′. $(X_1 + X_2) \cdot Y = X_1 \cdot Y + X_2 \cdot Y$.

La lunghezza è allora $\sqrt{X \cdot X}$ e il coseno dell'angolo fra due vettori non nulli è $(X \cdot Y)/(|X||Y|)$. Per giustificarlo occorre $(X \cdot Y)^2 \leqslant |X|^2 |Y|^2$, che si dimostra esattamente come prima: abbiamo usato solo gli assiomi 1′-4′, non la natura delle righe. Uno spazio lineare reale dotato di tale prodotto scalare si chiama spazio euclideo.

Nei diversi spazi concreti il prodotto scalare si introduce secondo le esigenze del problema. Per esempio, in opportuni spazi di funzioni reali $X(t)$ definite su $a \leqslant t \leqslant b$, si usano spesso $\int_a^b X(t)\, Y(t)\, dt$ oppure $\int_a^b X(t)\, Y(t)\, p(t)\, dt$, dove $p(t)$ è una funzione positiva. Per funzioni continue e, nel secondo caso, un peso positivo continuo, queste definizioni soddisfano gli assiomi 1′-4′.

Ortogonalità e basi ortonormali.

Due vettori sono ortogonali, o perpendicolari, se il loro prodotto scalare è zero. Vettori non nulli a due a due ortogonali sono indipendenti. Supponiamo infatti che X_1, ..., X_m abbiano queste proprietà e che $c_1X_1 + ... + c_mX_m = 0$.

Moltiplicando scalarmente per X_1 si ottiene $c_1|X_1|^2 = 0$, quindi $c_1 = 0$. Analogamente tutti gli altri coefficienti sono nulli: i vettori sono indipendenti.

In dimensione n non possono quindi esservi più di n vettori non nulli a due a due ortogonali; ogni famiglia di n tali vettori è una base. Se hanno tutti lunghezza 1, la base si dice ortonormale.

Si dimostra, ma omettiamo la prova, che ogni spazio euclideo ammette basi ortonormali, e in dimensione almeno due ne ammette infinite. Inoltre una base ortonormale di un sottospazio P può essere completata a una base ortonormale dell'intero spazio aggiungendo opportuni vettori.

È spesso conveniente usare le coordinate rispetto a una base ortonormale, perché il prodotto scalare assume una forma particolarmente semplice. Se X e Y hanno coordinate $(x_1, ..., x_n)$ e $(y_1, ..., y_n)$ nella base ortonormale $e_1, ..., e_n$, cioè

$$X = x_1e_1 + ... + x_ne_n \quad \text{e} \quad Y = y_1e_1 + ... + y_ne_n.$$

dalle proprietà del prodotto scalare segue

$$\begin{gathered} X \cdot Y = x_1y_1e_1 \cdot e_1 + x_1y_2e_1 \cdot e_2 + ... + x_1y_ne_1 \cdot e_n \\ + x_2y_1e_2 \cdot e_1 + x_2y_2e_2 \cdot e_2 + ... + x_2y_ne_2 \cdot e_n \\ + ... \\ + x_ny_1e_n \cdot e_1 + ... + x_ny_ne_n \cdot e_n \\ = x_1y_1 + x_2y_2 + ... + x_ny_n, \end{gathered}$$

perché $e_i \cdot e_k = 0$ per $i \neq k$ e $e_i \cdot e_i = 1$. In particolare, $X \cdot X = x_1^2 + ... + x_n^2$.

In coordinate ortonormali, lunghezza e prodotto scalare si esprimono dunque con le stesse formule dello spazio delle righe.

Il passaggio da quel modello allo spazio euclideo assiomatico non introduce complicazioni, ma estende le applicazioni della teoria.

Consideriamo la proiezione ortogonale su un sottospazio. Sia P_m un sottospazio di dimensione m dello spazio euclideo R_n. Sia $e_1, ..., e_m, f_1, ..., f_{n-m}$ una base ortonormale di R_n

che completa una base ortonormale di P_m. Il sottospazio Q_{n-m} generato dagli f_j si chiama complemento ortogonale di P_m e ha dimensione $n - m$. Esso consiste esattamente dei vettori ortogonali a tutti i vettori di P_m.

Ogni Z di R_n si scrive in modo unico come somma di un vettore X di P_m e uno Y di Q_{n-m}. Infatti la sua espressione nella base è unica:

$$Z = x_1 e_1 + \cdots + x_m e_m + y_1 f_1 + \cdots + y_{n-m} f_{n-m} ,$$

e basta porre $X = x_1 e_1 + \ldots + x_m e_m$, $Y = y_1 f_1 + \ldots + y_{n-m} f_{n-m}$.

Il vettore X si chiama proiezione ortogonale di Z su P_m.

Spazi unitari.

Anche in uno spazio complesso si possono definire lunghezza e prodotto scalare. Si parte ancora dal prodotto scalare, associando a ogni coppia X, Y un numero complesso, non necessariamente reale, $X \cdot Y$. Valgono gli assiomi seguenti:

1″. $X \cdot X$ è reale e positivo se $X \neq 0$; $0 \cdot 0 = 0$.

2″. $Y \cdot X = (X \cdot Y)'$, dove l'apice indica il coniugato complesso.

3″. $(cX) \cdot Y = c(X \cdot Y)$ per ogni complesso c.

4″. $(X_1 + X_2) \cdot Y = X_1 \cdot Y + X_2 \cdot Y$.

Per le righe complesse si può definire il prodotto scalare di $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ come $x_1 y'_1 + \ldots + x_n y'_n$. Si verifica facilmente che gli assiomi sono soddisfatti.

La lunghezza è $\sqrt{X \cdot X}$. Qui non si introduce un concetto di angolo fra vettori complessi.

Uno spazio lineare complesso con un prodotto scalare che soddisfa gli assiomi 1″-4″ si chiama spazio unitario.

§3. Sistemi di equazioni lineari

Due equazioni in due incognite e tre equazioni in tre incognite.

Un sistema di due equazioni lineari in due incognite ha la forma generale

$$a_1x + b_1y = c_1\,,$$
$$a_2x + b_2y = c_2\,.$$

Moltiplichiamo la prima equazione per b_2, la seconda per $-b_1$ e sommiamo:

$$(a_1b_2 - a_2b_1)\,x = c_1b_2 - c_2b_1\,.$$

Analogamente, moltiplicando la prima per $-a_2$ e la seconda per a_1, otteniamo

$$(a_1b_2 - a_2b_1)\,y = a_1c_2 - a_2c_1\,.$$

Si ricavano facilmente x e y, purché $a_1b_2 - a_2b_1$ sia diverso da zero. Questa espressione si chiama determinante della matrice

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

dei coefficienti del sistema e si indica con

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

La definizione suggerisce lo schema di calcolo



,

che non richiede ulteriori spiegazioni.

Anche le espressioni $c_1b_2 - c_2b_1$ e $a_1c_2 - a_2c_1$ sono determinanti:

$$\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} \qquad \text{e} \qquad \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

Se dunque il determinante

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

non è nullo, si ottengono le formule risolutive

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}; \quad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}. \qquad (4)$$

A rigore il ragionamento non è ancora completo. Le operazioni effettuate mostrano che, se una soluzione esiste e il determinante non è nullo, essa deve essere data dalle (4). Occorre ancora verificare che i valori trovati soddisfino entrambe le equazioni del sistema; la verifica non presenta difficoltà.

Se il determinante dei coefficienti è diverso da zero, il sistema ha dunque un'unica soluzione, data dalle (4).

Per un sistema di tre equazioni in tre incognite,

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1\,, \\ a_2x + b_2y + c_2z &= d_2\,, \\ a_3x + b_3y + c_3z &= d_3\,, \end{aligned}$$

si ragiona analogamente: si moltiplicano le equazioni per fattori opportuni e si sommano in modo da eliminare due incognite. Per eliminare y e z, i fattori sono $b_2c_3 - b_3c_2$, $b_3c_1 - b_1c_3$ e $b_1c_2 - b_2c_1$, come si verifica direttamente.

Il risultato è che, se l'espressione

$$\Delta = a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1$$

non è nulla, il sistema ha un'unica soluzione data da

$$x = \frac{\Delta_1}{\Delta}, \quad y = \frac{\Delta_2}{\Delta}, \quad z = \frac{\Delta_3}{\Delta},$$

dove $\Delta_1, \Delta_2, \Delta_3$ si ottengono da Δ sostituendo i coefficienti della corrispondente incognita con i termini noti.

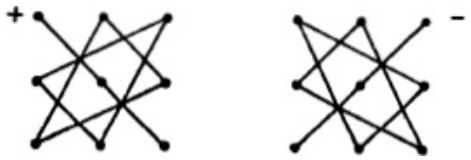
L'espressione Δ è il determinante della matrice

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

e si indica con

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Per calcolarlo è utile lo schema seguente:



Nel primo schema, le linee, una diagonale e due triangoli, congiungono gli elementi i cui prodotti compaiono con segno positivo; nel secondo congiungono quelli i cui prodotti compaiono con segno negativo.

Per due e per tre equazioni abbiamo ottenuto risultati analoghi: se il determinante dei coefficienti non è nullo, esiste un'unica soluzione. Ogni incognita è il rapporto fra due determinanti: al denominatore quello dei coefficienti; al numeratore quello ottenuto sostituendo la colonna dell'incognita con i termini noti.

Estendere direttamente il procedimento a n equazioni in n incognite è meno agevole. È più semplice generalizzare prima il determinante alle matrici quadrate di ordine arbitrario, studiarne le proprietà e applicarle poi ai sistemi.

Determinanti di ordine n.

Esaminiamo le espressioni esplicite

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1$$

e

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1 ,$$

In ogni termine compare esattamente un elemento per riga e uno per colonna, e tutti i prodotti possibili di questo tipo compaiono con un segno positivo o negativo. Questa osservazione suggerisce la definizione generale: il determinante di una matrice quadrata di ordine n è la somma algebrica di tutti i prodotti ottenuti scegliendo un elemento per ogni riga e ogni colonna, con segni assegnati secondo una regola precisa. Non ci soffermiamo su questa regola; basta sapere che garantisce la proprietà fondamentale seguente.

1. Scambiando due righe, il determinante cambia segno.

Per gli ordini 2 e 3 si verifica direttamente; nel caso generale la dimostrazione usa la regola dei segni qui non formulata.

I determinanti hanno molte altre proprietà che ne permettono l'uso in calcoli teorici e numerici, nonostante la complessità della loro espressione esplicita: un determinante di ordine n contiene $n!$ termini, ciascuno prodotto di n fattori, con segni stabiliti dalla regola ricordata.

Elenchiamo le proprietà principali, omettendo le dimostrazioni dettagliate. La prima è già stata enunciata.

2. Il determinante non cambia trasponendo la matrice, cioè scambiando righe e colonne senza alterarne l'ordine. La prova richiede lo studio della regola dei segni. Questa proprietà permette di trasferire alle colonne ogni affermazione sulle righe.

3. Il determinante è una funzione lineare omogenea degli elementi di ciascuna riga o colonna, quando le altre rimangono fisse. Precisamente,

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{i1} & \cdots & a_{in} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}\,, \qquad (5)$$

dove A_{i1}, ..., A_{in} non dipendono dagli elementi della riga i.

Infatti ogni termine contiene uno e un solo fattore proveniente da quella riga.

La (5) si chiama sviluppo del determinante secondo la riga i; i coefficienti A_{i1}, ..., A_{in} sono i complementi algebrici, o cofattori, degli elementi corrispondenti.

4. Il complemento algebrico A_{ij} coincide, a meno del segno, con il minore Δ_{ij}, determinante di ordine $n - 1$ ottenuto cancellando la riga i e la colonna j. Precisamente, il segno è $(-1)^{i+j}$. Le proprietà 3 e 4 riducono il calcolo di un determinante di ordine n a quello di n determinanti di ordine $n - 1$.

Dalle proprietà fondamentali seguono varie conseguenze; ne ricordiamo alcune.

5. Un determinante con due righe uguali è nullo.

Scambiando quelle righe il determinante non cambia, perché sono identiche, ma per la proprietà 1 deve cambiare segno: è dunque zero.

6. La somma dei prodotti degli elementi di una riga per i complementi algebrici corrispondenti di un'altra riga è zero.

Questa somma è infatti lo sviluppo, secondo una delle righe uguali, di un determinante con due righe uguali.

7. Un fattore comune agli elementi di una riga si può portare fuori dal segno di determinante.

È una conseguenza della proprietà 3.

8. Un determinante con due righe proporzionali è nullo.

Basta estrarre il fattore di proporzionalità: rimane un determinante con due righe uguali.

9. Il determinante non cambia aggiungendo a una riga un multiplo di un'altra.

Per la proprietà 3, il determinante modificato è infatti la somma di quello iniziale e di un determinante con due righe proporzionali, quindi nullo.

Quest'ultima proprietà offre un buon metodo di calcolo. Con tali operazioni si può annullare ogni elemento di una riga o colonna tranne uno, senza modificare il valore del determinante. Lo sviluppo lungo quella riga o colonna riduce allora il calcolo a un solo determinante di ordine $n - 1$, moltiplicato per l'elemento rimasto e per il segno del suo cofattore.

Calcoliamo, per esempio,

$$\Delta = \begin{vmatrix} 1 & 1 & -1 & 2 \\ 2 & -1 & 1 & 1 \\ -1 & 2 & 0 & 1 \\ 1 & 1 & -2 & 1 \end{vmatrix}.$$

Alla seconda colonna sottraiamo la prima; alla terza aggiungiamo la prima; alla quarta sottraiamo il doppio della prima. Otteniamo

$$\Delta = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -3 & 3 & -3 \\ -1 & 3 & -1 & 3 \\ 1 & 0 & -1 & -1 \end{vmatrix}.$$

Sviluppando lungo la prima riga,

$$\Delta = 1 \cdot (-1)^{1+1} \begin{vmatrix} -3 & 3 & -3 \\ 3 & -1 & 3 \\ 0 & -1 & -1 \end{vmatrix}.$$

Infine sommiamo la seconda riga alla prima e sviluppiamo lungo la prima colonna:

$$\Delta = \begin{vmatrix} 0 & 2 & 0 \\ 3 & -1 & 3 \\ 0 & -1 & -1 \end{vmatrix} = 3 \cdot (-1)^{1+2} \begin{vmatrix} 2 & 0 \\ -1 & -1 \end{vmatrix} = -3 \cdot (-2) = 6.$$

Il determinante di A si indica con $|A|$.

Concludiamo con un'altra proprietà fondamentale.

Il determinante del prodotto di due matrici quadrate dello stesso ordine è il prodotto dei loro determinanti: $|AB| = |A|\ |B|$.

Si possono quindi moltiplicare determinanti dello stesso ordine usando la regola del prodotto delle matrici.

Sistemi di *n* equazioni in *n* incognite.

I determinanti consentono di estendere facilmente i risultati ottenuti per due e tre incognite, supponendo che il determinante della matrice dei coefficienti sia diverso da zero.

Sia dato il sistema

$$\begin{aligned} &a_{11}x_1 + \cdots + a_{1j}x_j + \cdots + a_{1n}x_n = b_1 , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &a_{n1}x_1 + \cdots + a_{nj}x_j + \cdots + a_{nn}x_n = b_n \end{aligned}$$

Indichiamo con Δ il determinante dei coefficienti, non nullo per ipotesi, e con A_{ij} il complemento algebrico di a_{ij}. Moltiplichiamo la prima equazione per A_{1j}, la seconda per A_{2j}, e così via, e sommiamo:

$$\Delta x_j = b_1 A_{1j} + \cdots + b_n A_{nj} .$$

I coefficienti di tutte le incognite, salvo x_j, si annullano per la proprietà 6 applicata alle colonne. Quello di x_j è invece lo sviluppo del determinante lungo la colonna j, dunque è Δ.

Pertanto

$$x_j = (b_1 A_{1j} + \ldots + b_n A_{nj})/\Delta, \quad \text{per } j = 1, 2, \ldots, n. \quad (6)$$

Come già osservato, il ragionamento presuppone che $x_1, \ldots, x_n$ costituiscano una soluzione del sistema.

La conclusione è perciò questa:

se una soluzione esiste, è data dalle formule (6) ed è unica.

Per dimostrarne l'esistenza si sostituiscono i valori trovati in tutte le equazioni. Usando le stesse proprietà dei determinanti, questa volta applicate alle righe, si verifica che tutte sono soddisfatte.

Vale dunque il teorema: se il determinante dei coefficienti di un sistema di n equazioni in n incognite è diverso da zero, il sistema ha un'unica soluzione, data dalle (6).

Le formule si possono riscrivere osservando che $b_1 A_{1j} + \ldots + b_n A_{nj}$ è il determinante

$$\Delta_j = b_1 A_{1j} + \cdots + b_n A_{nj} = \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}$$

in cui i termini noti occupano la colonna j.

I risultati per due e tre incognite sono così generalizzati completamente: anche la forma delle formule risolutive rimane identica.

Un corollario: se un sistema quadrato non ha soluzioni, oppure ne ha più di una, il determinante dei coefficienti è nullo.

Il corollario si applica spesso ai sistemi omogenei, nei quali tutti i termini noti sono zero. Essi ammettono sempre la soluzione banale $x_1 = \ldots = x_n = 0$.

Se un sistema omogeneo quadrato ammette anche una soluzione non banale, il suo determinante è nullo.

Questa osservazione permette numerose applicazioni dei determinanti in altri rami della matematica.

Consideriamo il problema geometrico di trovare il piano passante per tre punti non allineati (x_1, y_1, z_1), (x_2, y_2, z_2) e (x_3, y_3, z_3).

La geometria elementare garantisce che esiste. Se la sua equazione è $Ax + By + Cz + D = 0$, allora

$$\begin{aligned} Ax_1 + By_1 + Cz_1 + D &= 0, \\ Ax_2 + By_2 + Cz_2 + D &= 0, \\ Ax_3 + By_3 + Cz_3 + D &= 0. \end{aligned}$$

Per un punto arbitrario (x, y, z) del piano vale inoltre

$$Ax + By + Cz + D = 0.$$

Consideriamo le quattro equazioni come un sistema omogeneo nelle incognite A, B, C, D. Poiché il piano esiste, il sistema ha una soluzione non banale; il suo determinante è dunque zero:

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x & y & z & 1 \end{vmatrix} = 0. \qquad (7)$$

Questa è l'equazione cercata. È di primo grado in x, y, z, per la linearità del determinante rispetto all'ultima riga.

Poiché i tre punti non sono allineati, si verifica che i coefficienti dell'equazione non sono tutti nulli. La (7) rappresenta quindi un piano, e passa per i punti dati perché le loro coordinate la soddisfano.

Scrittura matriciale di un sistema quadrato.

Un sistema di n equazioni lineari in n incognite,

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

si scrive come un'unica equazione matriciale:

$$AX = B.$$

Qui A è la matrice dei coefficienti, X la colonna delle incognite e B quella dei termini noti.

Se il determinante di A è non nullo, la soluzione si scrive esplicitamente, secondo la (6),

$$x_1 = \frac{A_{11}}{\Delta} b_1 + \frac{A_{21}}{\Delta} b_2 + \cdots + \frac{A_{n1}}{\Delta} b_n ,$$

$$x_2 = \frac{A_{12}}{\Delta} b_1 + \frac{A_{22}}{\Delta} b_2 + \cdots + \frac{A_{n2}}{\Delta} b_n ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$x_n = \frac{A_{1n}}{\Delta} b_1 + \frac{A_{2n}}{\Delta} b_2 + \cdots + \frac{A_{nn}}{\Delta} b_n$$

oppure, in forma matriciale,

$$X = \begin{bmatrix} \frac{A_{11}}{\Delta} & \frac{A_{21}}{\Delta} & \cdots & \frac{A_{n1}}{\Delta} \\ \frac{A_{12}}{\Delta} & \frac{A_{22}}{\Delta} & \cdots & \frac{A_{n2}}{\Delta} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{A_{1n}}{\Delta} & \frac{A_{2n}}{\Delta} & \cdots & \frac{A_{nn}}{\Delta} \end{bmatrix} B.$$

Il primo fattore a destra si chiama matrice inversa di A e si indica con A^{-1}. La soluzione di $AX = B$ assume allora la forma semplice e naturale

$$X = A^{-1}B.$$

Possiamo dimostrare il risultato anche usando direttamente l'algebra delle matrici.

Ricordiamo anzitutto il ruolo particolare della matrice

$$E = \begin{pmatrix} 1 & 0 & 0 \cdots 0 \\ 0 & 1 & 0 \cdots 0 \\ \cdots & \cdots & \cdots \\ 0 & 0 & 0 \cdots 1 \end{pmatrix},$$

detta matrice identità, o unitaria nell'originale.

Fra le matrici quadrate essa svolge il ruolo del numero 1: $AE = EA = A$. La verifica segue immediatamente dalla regola del prodotto.

Analogamente, la matrice inversa svolge rispetto ad A il ruolo del reciproco di un numero:

$$AA^{-1} = A^{-1}A = E.$$

Le uguaglianze si verificano mediante il prodotto matriciale e le proprietà 3 e 6 dei determinanti.

Possiamo quindi risolvere $AX = B$ come segue.

Se $AX = B$, allora $A^{-1}(AX) = A^{-1}B$. Ma $A^{-1}(AX) = (A^{-1}A)X = EX = X$, dunque $X = A^{-1}B$.

Viceversa, se $X = A^{-1}B$, allora $AX = AA^{-1}B = EB = B$.

L'equazione matriciale ammette quindi l'unica soluzione $X = A^{-1}B$, purché l'inversa esista.

Abbiamo dimostrato che l'inversa esiste se il determinante è non nullo. La condizione è anche necessaria: se $AA^{-1} = E$, prendendo i determinanti si ha

$$| A | | A^{-1} | = | E | = 1,$$

e quindi il determinante di A non è zero.

Una matrice con determinante non nullo si chiama non degenere, o non singolare. Le matrici invertibili sono dunque esattamente quelle non singolari.

Il concetto di inversa è utile non soltanto per i sistemi, ma anche in molti altri problemi dell'algebra lineare.

Le formule risolutive mediante determinanti sono preziose nella teoria, ma poco convenienti per il calcolo numerico. Per quest'ultimo sono stati sviluppati metodi e schemi diversi; data l'importanza pratica dei sistemi, soprattutto con molte incognite, la ricerca di procedimenti più efficienti è intensa nell'epoca descritta dall'autore.

Il caso generale dei sistemi lineari.

Non supponiamo più uguali il numero delle equazioni e quello delle incognite. Non possiamo aspettarci sempre l'esistenza né l'unicità della soluzione. Se le equazioni sono meno delle incognite, ci si attende una molteplicità di soluzioni: due equazioni in tre incognite, per esempio, sono soddisfatte dai punti della retta d'intersezione dei piani corrispondenti. Ma possono anche non esserci soluzioni, se i piani sono paralleli e distinti. Se le equazioni sono più numerose delle incognite, in generale non vi sono soluzioni; tuttavia possono essercene, anche infinite.

Per studiare l'esistenza e la struttura delle soluzioni ricorreremo a un'interpretazione geometrica.

Interpretiamo il sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \cdots\cdots\cdots\

L'esistenza di una soluzione equivale a chiedere se B sia combinazione lineare delle colonne $A_1, \ldots, A_n$.

Ciò accade se e solo se B appartiene al sottospazio da esse generato, ossia se le famiglie $A_1, \ldots, A_n$ e $A_1, \ldots, A_n, B$ generano lo stesso sottospazio.

Poiché il primo sottospazio è contenuto nel secondo, essi coincidono se e solo se hanno uguale dimensione. Ricordando la definizione di rango, il sistema $x_1A_1 + \ldots + x_nA_n = B$ è dunque risolubile esattamente quando le due famiglie hanno lo stesso rango.

Si dimostra, ma non lo faremo qui, che il rango di una famiglia di vettori coincide con quello della matrice delle loro coordinate. Il rango di una matrice si può definire come il massimo ordine di un suo minore non nullo, ottenuto eliminando alcune righe e colonne; alla matrice nulla si attribuisce rango zero.

Nella base naturale delle colonne, le coordinate degli A_j sono i coefficienti del sistema e quelle di B sono i termini noti. Otteniamo così il criterio definitivo.

Il sistema lineare

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

ammette almeno una soluzione se e solo se la matrice dei coefficienti e la matrice completa, ottenuta aggiungendo la colonna dei termini noti, hanno lo stesso rango.

Studiamo la struttura delle soluzioni quando esistono. Sia $x_1^0, x_2^0, \cdots, x_n^0$ una soluzione della (8). Poniamo $x_1 = x_1^0 + y_1$, $x_2 = x_2^0 + y_2, \cdots, x_n = x_n^0 + y_n$. Poiché $x_1^0, x_2^0, \cdots, x_n^0$ risolve la (8), le nuove incognite $y_1, \ldots, y_n$ soddisfano il sistema omogeneo

$$\begin{aligned} a_{11}y_1 + \cdots + a_{1n}y_n &= 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ a_{m1}y_1 + \cdots + a_{mn}y_n &= 0 \end{aligned} \qquad (9)$$

con la stessa matrice dei coefficienti. Viceversa, aggiungendo alla soluzione particolare $x_1^0, x_2^0, \cdots, x_n^0$ una qualunque soluzione del sistema omogeneo (9), si ottiene una soluzione della (8).

La soluzione generale della (8) è dunque la somma di una sua soluzione particolare e della soluzione generale del sistema omogeneo associato.

Basta quindi studiare l'insieme delle soluzioni della (9), cosa che faremo ora.

Sistemi omogenei.

Interpretiamo il sistema omogeneo

$$\begin{aligned} a_{11}y_1 &+ a_{12}y_2 + \cdots + a_{1n}y_n = 0, \\ a_{21}y_1 &+ a_{22}y_2 + \cdots + a_{2n}y_n = 0, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}y_1 &+ a_{m2}y_2 + \cdots + a_{mn}y_n = 0 \end{aligned}$$

nello spazio euclideo a n dimensioni. Supponiamo reali i coefficienti; nel caso complesso si procede analogamente nello spazio unitario, con risultati corrispondenti.

Siano $A_1', A_2', \cdots, A_m'$ e Y i vettori le cui coordinate in una base ortonormale sono rispettivamente

$$(a_{11}, a_{12}, \cdots, a_{1n}),\ (a_{21}, a_{22}, \cdots, a_{2n}),\ \cdots,\ (a_{m1}, a_{m2}, \cdots, a_{mn}),$$
$$(y_1, y_2, \cdots, y_n).$$

Il sistema assume allora la forma

$$A_1'Y = 0,\ A_2'Y = 0,\ \cdots,\ A_m'Y = 0;$$

cioè ogni soluzione è un vettore ortogonale a tutti i vettori formati dai coefficienti delle singole equazioni.

Le soluzioni formano quindi il complemento ortogonale del sottospazio generato da $A_1', A_2', \cdots, A_m'$. Quest'ultimo ha dimensione pari al rango r della matrice dei coefficienti; lo spazio delle soluzioni ha pertanto dimensione $n - r$.

Ogni sottospazio ha una base. Esistono dunque $n - r$ soluzioni indipendenti del sistema omogeneo, le cui combinazioni lineari danno tutte le soluzioni; qui n è il numero delle incognite e r il rango dei coefficienti.

La struttura delle soluzioni dei sistemi omogenei e non omogenei è così chiarita. Un sistema omogeneo ammette soltanto la soluzione nulla se e solo se il rango è uguale al numero delle incognite. La stessa condizione garantisce l'unicità per un sistema non omogeneo, purché sia soddisfatta la condizione di compatibilità.

Questo studio mostra come i concetti geometrici generalizzati rendano semplice e chiaro un problema algebrico complesso.

§4. Trasformazioni lineari

Definizione ed esempi.

In molte ricerche occorre cambiare variabili, passando da $x_1, \ldots, x_n$ a $y_1, \ldots, y_n$, legate alle prime da relazioni funzionali:

$$\begin{aligned} y_1 &= \phi_1(x_1, x_2, \cdots, x_n), \\ y_2 &= \phi_2(x_1, x_2, \cdots, x_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= \phi_n(x_1, x_2, \cdots, x_n). \end{aligned}$$

Per esempio, un cambiamento di coordinate nel piano o nello spazio esprime le vecchie coordinate in funzione delle nuove, o viceversa.

Una trasformazione di variabili compare anche quando cambia la posizione o configurazione di un oggetto descritto da quelle variabili. Un esempio tipico è la variazione delle coordinate dei punti di un corpo durante una deformazione.

Una trasformazione astratta di n variabili si interpreta usualmente come una trasformazione dello spazio a n dimensioni: a ogni vettore del dominio, di coordinate $x_1, \ldots, x_n$, associa un vettore di coordinate $y_1, \ldots, y_n$.

Poiché una funzione con derivate parziali continue è localmente approssimata da una funzione lineare, ogni trasformazione liscia, le cui funzioni $\phi_1, \phi_2, \cdots, \phi_n$ abbiano derivate parziali continue, è localmente vicina a una trasformazione della forma

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n. \end{aligned} \qquad (10)$$

Già questo rende fondamentale lo studio delle trasformazioni lineari. Dalla teoria dei sistemi sappiamo che la (10) si risolve univocamente rispetto alle x_j, cioè la trasformazione è invertibile, se e solo se il determinante dei coefficienti è non nullo. Su questo fatto si fonda un importante teorema dell'analisi: una trasformazione

$$\begin{aligned} y_1 &= \phi_1(x_1, x_2, \cdots, x_n), \\ y_2 &= \phi_2(x_1, x_2, \cdots, x_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= \phi_n(x_1, x_2, \cdots, x_n), \end{aligned}$$

continuamente differenziabile vicino a un punto ammette localmente un'inversa continuamente differenziabile se e solo se, in quel punto, il determinante

$$\begin{vmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \frac{\partial \phi_2}{\partial x_1} & \cdots & \frac{\partial \phi_2}{\partial x_n} \\ \cdots\cdots & & \\ \frac{\partial \phi_n}{\partial x_1} & \cdots & \frac{\partial \phi_n}{\partial x_n} \end{vmatrix}$$

è diverso da zero.

Lo studio della trasformazione generale (10) si riduce essenzialmente a quello della trasformazione omogenea con gli stessi coefficienti,

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n , \\ y_2 &= a_{21}x_1 + \cdots + a_{2n}x_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + \cdots + a_{nn}x_n , \end{aligned} \qquad (11)$$

e d'ora in poi per trasformazione lineare intenderemo sempre una trasformazione omogenea.

Le trasformazioni lineari si possono definire anche attraverso proprietà intrinseche, senza usare le formule in coordinate (11). Questa definizione ha il vantaggio di non dipendere dalla base.

Una trasformazione lineare di uno spazio è una funzione $Y = A(X)$, fra vettori, che soddisfa il postulato di linearità

$$A(c_1X_1 + c_2X_2) = c_1A(X_1) + c_2A(X_2). \qquad (12)$$

Nel seguito useremo questa definizione.

Essa equivale a quella in coordinate. Una funzione che esprime le coordinate di Y come forme lineari delle coordinate di X soddisfa evidentemente la (12). Viceversa, se A soddisfa la (12) e $e_1, \ldots, e_n$ è una base, allora

$$A(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1A(e_1) + x_2A(e_2) + \cdots + x_nA(e_n).$$

Indichiamo con $a_{1j}, \ldots, a_{nj}$ le coordinate di $A(e_j)$ nella stessa base. Le coordinate di $Y = A(X)$ sono

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n , \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n . \end{aligned}$$

A ogni trasformazione lineare corrisponde quindi, scelta una base, una matrice quadrata. Scriviamo $Y = AX$, dove X e Y sono le colonne delle coordinate del vettore iniziale e della sua

immagine. Le colonne di A sono le coordinate delle immagini dei vettori della base. Seguendo questa notazione, scriveremo spesso anche la trasformazione stessa come $Y = AX$, senza parentesi.

Dalla formula

$$A(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1A(e_1) + x_2A(e_2) + \cdots + x_nA(e_n)$$

segue che l'immagine dell'intero spazio è il sottospazio generato da $A(e_1)$, ..., $A(e_n)$. La sua dimensione è il rango di questa famiglia, quindi della matrice A. L'immagine coincide con l'intero spazio se e solo se il rango è n, ossia il determinante è non nullo. La trasformazione si dice allora non singolare, o non degenere.

Essa è invertibile e le coordinate del vettore iniziale si ricavano mediante $X = A^{-1}Y$.

Una trasformazione con determinante nullo si dice singolare, o degenere. Non è invertibile: la sua immagine è un sottospazio proprio dello spazio ambiente.

Un primo esempio non degenere è l'identità, che lascia fisso ogni vettore: in ogni base la sua matrice è E. Anche l'omotetia che moltiplica ogni vettore per uno stesso numero non nullo a è non degenere. La sua matrice è aE, indipendentemente dalla base.

Un caso importante è quello delle trasformazioni ortogonali: trasformazioni lineari di uno spazio euclideo che conservano le lunghezze. Esse generalizzano in dimensione n le rotazioni dello spazio tridimensionale intorno all'origine e le loro composizioni con una riflessione rispetto a un piano passante per l'origine.

Conservano anche i prodotti scalari e quindi trasformano una base ortonormale in una famiglia di vettori unitari a due a due ortogonali, che è ancora una base.

La loro matrice rispetto a una base ortonormale ha proprietà specifiche.

La somma dei quadrati degli elementi di ciascuna colonna è 1, perché è il quadrato della lunghezza dell'immagine di un vettore della base. La somma dei prodotti degli elementi corrispondenti di due colonne diverse è zero, perché è il prodotto scalare di due immagini ortogonali.

In forma matriciale, entrambe le proprietà si esprimono con

$$\bar{P}P = E.$$

dove P è la matrice della trasformazione e $\bar{P}$ la sua trasposta, le cui righe sono le colonne di P.

Infatti gli elementi diagonali di $\bar{P}P$ sono le somme dei quadrati nelle colonne di P; quelli fuori diagonale sono le somme dei prodotti fra elementi corrispondenti di colonne diverse.

Un esempio degenere è la proiezione ortogonale su un sottospazio proprio: l'intero spazio viene mandato in una sua parte.

Cambiamento di coordinate.

Studiamo come cambiano le coordinate passando da una base a un'altra.

Siano $e_1, \ldots, e_n$ la vecchia base e $f_1, \ldots, f_n$ la nuova. Sia

$$C = \begin{bmatrix} c_{11}, \cdots, c_{1n} \\ \cdots\cdots\cdots \\ c_{n1}, \cdots, c_{nn} \end{bmatrix}$$

la matrice che ha per colonne le coordinate degli f_j nella vecchia base. Essa è invertibile, poiché gli f_j sono indipendenti, e si chiama matrice del cambiamento di coordinate.

Indichiamo con $x_1, \ldots, x_n$ e $x'_1, \ldots, x'_n$ le vecchie e le nuove coordinate dello stesso vettore X. Allora $X = x'_1 f_1 + x'_2 f_2 + \cdots + x'_n f_n$ e le vecchie coordinate formano la colonna

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \begin{bmatrix} c_{11}x'_1 + c_{12}x'_2 + \cdots + c_{1n}x'_n \\ c_{21}x'_1 + c_{22}x'_2 + \cdots + c_{2n}x'_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_{n1}x'_1 + c_{n2}x'_2 + \cdots + c_{nn}x'_n \end{bmatrix} = \begin{bmatrix} c_{11} \cdots c_{1n} \\ c_{21} \cdots c_{2n} \\ \cdots\cdots\cdots \\ c_{n1} \cdots c_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ \cdot \\ \cdot \\ x'_n \end{bmatrix}.$$

Esse si esprimono dunque linearmente e omogeneamente nelle nuove mediante la matrice C.

Queste formule coincidono formalmente con quelle che collegano le coordinate di un vettore e della sua immagine sotto una trasformazione lineare invertibile. Una stessa sostituzione lineare omogenea può quindi essere interpretata come cambiamento di coordinate oppure come trasformazione dello spazio; l'interpretazione dipende dal problema concreto.

Vediamo ora come cambia la matrice di una trasformazione quando cambiamo base.

Nella vecchia base la trasformazione abbia matrice A:

$$Y = AX.$$

Se X' e Y' sono le coordinate nella nuova base, allora $X = CX'$ e $Y = CY'$, perciò

$$Y' = C^{-1}Y = C^{-1}AX = C^{-1}ACX'.$$

La nuova matrice è dunque $C^{-1}AC$.

Due matrici legate da $B = C^{-1}AC$, con C invertibile, si dicono simili. Una stessa trasformazione corrisponde, nelle diverse basi, a matrici tutte simili fra loro.

Autovettori e autovalori.

Consideriamo una classe importante di trasformazioni costruite nel modo seguente.

Siano e_1, ..., e_n vettori indipendenti che vengono moltiplicati rispettivamente per λ_1, ..., λ_n. Prendendoli come base, la trasformazione si scrive

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Nello spazio reale di dimensione due o tre l'interpretazione è intuitiva. Se tutti i λ_i sono positivi, lo spazio viene dilatato o contratto nelle direzioni degli e_i con quei fattori. Un fattore negativo comporta anche l'inversione del verso corrispondente. Se, per esempio, $\lambda_1 = 0$, si ha una proiezione parallela a e_1 sul sottospazio generato dagli altri vettori, seguita dalle deformazioni nelle direzioni restanti.

Questa classe è importante perché, nonostante la semplicità, è molto generale. Sotto opportune condizioni una trasformazione ammette una base nella quale la sua matrice è diagonale.

Le condizioni si comprendono meglio nello spazio complesso, al quale ci riferiremo nel seguito.

Introduciamo una definizione.

Un vettore non nullo X si chiama autovettore di A se $AX = \lambda X$ per un numero λ, detto autovalore. La trasformazione manda dunque X in un suo multiplo, eventualmente nullo.

Se la matrice è diagonale in una base, i vettori della base sono autovettori e gli elementi diagonali sono i corrispondenti autovalori. Viceversa, una base di autovettori rende diagonale la matrice.

Studiamo questi concetti in coordinate. Siano A la matrice della trasformazione e X la colonna delle coordinate del vettore. L'equazione $AX = \lambda X$ equivale a

$$(A - \lambda E)\, X = 0.$$

ossia al sistema

$$\begin{aligned} (a_{11} - \lambda)\, x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0, \\ a_{21}x_1 + (a_{22} - \lambda)\, x_2 + \cdots + a_{2n}x_n &= 0, \\ \cdots\cdots\cdots\

È un sistema omogeneo nelle x_1, ..., x_n. Cerchiamo soluzioni non banali, perché un autovettore non è nullo. Il rango deve quindi essere minore di n, condizione equivalente a

$$\begin{vmatrix} a_{11}-\lambda & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0.$$

Gli autovalori sono dunque esattamente le radici di $|A - \lambda E|$; a ogni radice corrisponde almeno un autovettore. Questo polinomio è il polinomio caratteristico di A; l'equazione che lo annulla è detta caratteristica, o secolare, e le radici sono anche chiamate numeri caratteristici.[1]

Per il teorema fondamentale dell'algebra, un polinomio non costante ha almeno una radice complessa. Ogni trasformazione di uno spazio complesso di dimensione positiva ammette quindi autovalori e autovettori. Anche per una matrice reale, però, alcuni o tutti gli autovalori possono essere non reali: il teorema non garantisce autovettori reali. Per esempio, una rotazione del piano di angolo non multiplo di 180° non lascia invariata alcuna direzione reale, e non ha autovettori reali.

Matrici che rappresentano la stessa trasformazione hanno gli stessi autovalori. È plausibile che abbiano anche lo stesso polinomio caratteristico, indipendentemente dalla base. Lo verifichiamo con un breve calcolo.

La nuova matrice è $C^{-1}AC$, con C invertibile. Ma

$$\begin{aligned} |\, C^{-1}AC - \lambda E \,| &= |\, C^{-1}AC - C^{-1}\lambda EC \,| = |\, C^{-1}(A - \lambda E)\, C \,| \\ &= |\, C^{-1} \,|\, |\, C \,|\, |\, A - \lambda E \,| = |\, C^{-1}C \,|\, |\, A - \lambda E \,| = |\, A - \lambda E \,|. \end{aligned}$$

Il polinomio caratteristico è quindi lo stesso in ogni base e si può chiamare polinomio caratteristico della trasformazione.

Supponiamo che le n radici del polinomio caratteristico siano distinte. Dimostriamo che autovettori associati a valori diversi sono indipendenti. Altrimenti scegliamo una sottofamiglia indipendente massimale e_1, ..., e_k; un altro autovettore, e_{k+1}, si esprimerebbe come

$$e_{k+1} = c_1 e_1 + c_2 e_2 + \cdots + c_k e_k \,. \qquad (13)$$

Applicando la trasformazione,

$$Ae_{k+1} = c_1 Ae_1 + c_2 Ae_2 + \cdots + c_k Ae_k \,,$$

e quindi, per la proprietà degli autovettori,

$$\lambda_{k+1} e_{k+1} = c_1 \lambda_1 e_1 + c_2 \lambda_2 e_2 + \cdots + c_k \lambda_k e_k \,.$$

Moltiplichiamo la (13) per λ_{k+1} e sottraiamo l'ultima uguaglianza:

$$c_1(\lambda_{k+1} - \lambda_1)\, e_1 + c_2(\lambda_{k+1} - \lambda_2)\, e_2 + \cdots + c_k(\lambda_{k+1} - \lambda_k)\, e_k = 0,$$

Dall'indipendenza dei primi *k* vettori segue

$$c_1(\lambda_{k+1} - \lambda_1) = c_2(\lambda_{k+1} - \lambda_2) = \cdots = c_k(\lambda_{k+1} - \lambda_k) = 0.$$

Ma le differenze $\lambda_{k+1} - \lambda_j$ sono tutte non nulle. Tutti i coefficienti della (13) sarebbero allora nulli, mentre e_{k+1} non è nullo: una contraddizione.

Se il polinomio caratteristico ha *n* radici distinte, possiamo quindi scegliere un autovettore per ciascuna: sono *n* vettori indipendenti e formano una base. In essa la matrice è diagonale.

In linguaggio matriciale: una matrice con *n* autovalori distinti è simile alla matrice diagonale formata da quegli autovalori.

Se vi sono radici multiple, la riduzione alla forma più semplice è più complessa. Ci limiteremo a enunciare il risultato.

Un blocco canonico di ordine *m* è una matrice della forma

$$I_{m,\lambda_i} = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}.$$

in cui tutti gli elementi non indicati sono zero.

Una matrice canonica di Jordan ha blocchi canonici lungo la diagonale principale e zeri altrove:

$$\begin{bmatrix} I_{m_1,\lambda_1} & & & \\ & I_{m_2,\lambda_2} & & \\ & & \ddots & \\ & & & I_{m_k,\lambda_k} \end{bmatrix}.$$

Gli autovalori dei diversi blocchi possono coincidere. Ogni matrice complessa è simile a una matrice canonica di Jordan. La dimostrazione è piuttosto complessa; il risultato è fondamentale in molte applicazioni, in particolare nei sistemi di equazioni differenziali lineari.

Una matrice è diagonalizzabile se e solo se tutti i suoi blocchi di Jordan hanno ordine 1.

§5. Forme quadratiche

Definizione e prime proprietà.

Una forma quadratica è un polinomio omogeneo di grado due in più variabili.

Essa comprende quadrati delle variabili e prodotti di due variabili, moltiplicati per coefficienti. Si può disporre nello schema

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = \quad & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 . \end{aligned}$$

I termini simili, come $a_{12}x_1x_2$ e $a_{21}x_2x_1$, si scrivono con coefficienti uguali, ciascuno pari alla metà del coefficiente complessivo. Alla forma corrisponde così un'unica matrice simmetrica.

Indichiamo con X la colonna delle variabili e con $\tilde{X}$ la riga $(x_1, \ldots, x_n)$, sua trasposta. La notazione matriciale della forma è

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = \quad & x_1(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n) \\ & + x_2(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n) + \cdots \\ & + x_n(a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n) \end{aligned}$$

$$= (x_1, x_2, \cdots, x_n) \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{bmatrix}$$

$$= (x_1, x_2, \cdots, x_n) \begin{bmatrix} a_{11}a_{12} \cdots a_{1n} \\ a_{21}a_{22} \cdots a_{2n} \\ \cdots\cdots\cdots \\ a_{n1}a_{n2} \cdots a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \tilde{X}AX.$$

Le forme quadratiche compaiono in molti rami della matematica e nelle applicazioni.

In teoria dei numeri e cristallografia si considerano variabili intere; in geometria analitica compaiono nelle equazioni di coniche e quadriche; in meccanica e fisica descrivono, per esempio, l'energia cinetica in funzione delle velocità generalizzate. In analisi servono a studiare quanto una funzione si discosti, vicino a un punto, dalla propria approssimazione lineare: un esempio è la ricerca di massimi e minimi.

Sia $w = f(x, y)$ con derivate parziali continue fino all'ordine tre. Per un estremo in un punto interno (x_0, y_0) è necessario che le derivate prime si annullino. Supponiamolo e diamo incrementi h, k alle variabili. L'incremento $\Delta w = f(x_0 + h, y_0 + k) - f(x_0, y_0)$ coincide, a meno

di termini di ordine superiore, con $\frac{1}{2}(rh^2 + 2shk + tk^2)$, dove r, s, t sono le derivate seconde rispetto a x, a x e y, e a y, valutate nel punto. Se la forma è positiva per ogni (h, k) non nullo, vi è un minimo; se è negativa, un massimo; se assume entrambi i segni, non vi è né massimo né minimo. Analogamente si studiano funzioni di più variabili.

Lo studio delle forme quadratiche riguarda soprattutto l'equivalenza rispetto a determinate trasformazioni lineari delle variabili. Due forme sono equivalenti se una si ottiene dall'altra con una trasformazione ammessa. È strettamente connesso il problema della riduzione a una forma il più possibile semplice.

Le trasformazioni ammesse dipendono dal problema.

In analisi si ammettono trasformazioni lineari invertibili arbitrarie; in geometria interessano soprattutto quelle ortogonali, corrispondenti al cambio di assi cartesiani ortonormali. In teoria dei numeri e cristallografia si considerano trasformazioni a coefficienti interi con determinante 1.

Studieremo la riduzione mediante trasformazioni invertibili arbitrarie e mediante trasformazioni ortogonali. Vediamo prima come cambia la matrice della forma.

Sia $f(x_1, x_2, \cdots, x_n) = \bar{X}AX$, con A simmetrica e X colonna delle variabili.

Poniamo $X = CX'$. Allora $\bar{X} = \bar{X}'\bar{C}$ e quindi $\bar{X}AX = \bar{X}'(\bar{C}AC)\,X'$: la matrice trasformata è $\bar{C}AC$.

La matrice $\bar{C}AC$ è ancora simmetrica. Ridurre una forma quadratica equivale quindi a semplificare la sua matrice simmetrica moltiplicandola a destra per una matrice e a sinistra per la trasposta di questa.

Riduzione canonica mediante completamento dei quadrati.

Ogni forma quadratica reale si può trasformare, con una sostituzione lineare reale invertibile, in una somma di quadrati delle nuove variabili con opportuni coefficienti.

Mostriamo anzitutto che, per una forma non identicamente nulla, si può rendere non nullo il coefficiente del quadrato della prima variabile.

Sia infatti

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = \quad & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 . \end{aligned}$$

Se $a_{11} \neq 0$, non occorre fare nulla. Se è zero ma $a_{kk} \neq 0$ per qualche k, poniamo $x_1 = x_k'$, $x_k = x_1'$, lasciando invariate le altre variabili: la trasformazione è invertibile e

raggiunge lo scopo. Se invece tutti gli elementi diagonali sono nulli, almeno uno fuori diagonale è non nullo, per esempio a_{12}.

Usiamo allora la trasformazione invertibile

$$x_1 = x'_1,$$
$$x_2 = x'_1 + x'_2.$$

e lasciamo invariate le altre variabili: il risultato desiderato è ottenuto.

Possiamo dunque supporre senza perdita di generalità $a_{11} \neq 0$.

Isoliamo ora il quadrato di una forma lineare che raccolga tutti i termini contenenti x_1.

È facile farlo, perché

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) = {} & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \end{aligned}$$

$$\begin{aligned} = {} & a_{11}\left[x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right]^2 - a_{11}\left[\frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right]^2 \\ & + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2. \end{aligned}$$

Sviluppando il secondo addendo e raccogliendo i termini simili, si ottiene

$$f(x_1, x_2, \cdots, x_n) = a_{11}\left[x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right]^2 + f_1(x_2, \cdots, x_n),$$

dove f_1 dipende da $n - 1$ variabili.

La trasformazione

$$\begin{aligned} x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n &= x'_1, \\ x_2 &= x'_2, \\ \cdots\cdots\cdots\cdots & \\ x_n &= x'_n, \end{aligned}$$

è evidentemente non singolare. Eseguendo questa trasformazione, riduciamo la forma a

$$a_{11}x_1'^2 + f_1(x'_2, \cdots, x'_n).$$

Continuando nello stesso modo, riduciamo la forma alla «forma canonica» richiesta:

$$\alpha_1 z_1^2 + \alpha_2 z_2^2 + \cdots + \alpha_n z_n^2 .$$

Qui z_1, z_2, ..., z_n sono le nuove variabili introdotte nell'ultimo passaggio.

La legge d'inerzia delle forme quadratiche.

Nella riduzione di una forma quadratica alla forma canonica resta sempre una notevole arbitrarietà nella scelta della trasformazione delle variabili. Essa deriva, fra l'altro, dalla possibilità di far precedere il metodo di separazione successiva dei quadrati da una trasformazione non singolare arbitraria delle variabili.

Nonostante questa arbitrarietà, la riduzione conduce a una forma quadratica canonica quasi unica, indipendente dalla trasformazione scelta. Precisamente, il numero dei quadrati delle nuove variabili che compaiono con coefficienti positivi, negativi e nulli è sempre lo stesso, qualunque sia il metodo di riduzione. Questo teorema è noto come legge d'inerzia delle forme quadratiche. Non lo dimostreremo qui.

La legge d'inerzia risolve il problema dell'equivalenza delle forme quadratiche reali rispetto a tutte le trasformazioni non singolari. Due forme sono infatti equivalenti se e solo se le rispettive riduzioni canoniche presentano lo stesso numero di quadrati con coefficienti positivi, negativi e nulli.

Particolarmente importanti nelle applicazioni sono le forme quadratiche la cui forma canonica è una somma di quadrati delle nuove variabili con coefficienti tutti positivi. Queste forme si chiamano definite positive.

Le forme quadratiche definite positive sono caratterizzate dalla proprietà di assumere valori positivi ogni volta che alle variabili si assegnano valori reali non tutti nulli.

Riduzione ortogonale delle forme quadratiche alla forma canonica.

Fra tutti i metodi di riduzione alla forma canonica rivestono particolare interesse le trasformazioni ortogonali, ossia quelle ottenute mediante una trasformazione lineare delle variabili con matrice ortogonale. Esse intervengono, per esempio, in geometria analitica, quando si riduce alla forma canonica l'equazione generale di una curva o di una superficie del secondo ordine.

Per dimostrare che una simile riduzione è possibile, conviene considerare la forma quadratica come funzione dei vettori di uno spazio euclideo, interpretando le variabili x_1, x_2, ..., x_n come coordinate di un vettore variabile in una base ortonormale. Una trasformazione ortogonale delle variabili corrisponde allora al passaggio da una base ortonormale a un'altra.

Alla forma quadratica

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + \cdots + a_{1n}x_1x_n$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$+ a_{n1}x_nx_1 + \cdots + a_{nn}x_n^2$$

associamo la trasformazione lineare A che, nella base scelta, ha matrice

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

La nostra forma quadratica si può allora interpretare come il prodotto scalare $AX \cdot X$, dove X ha coordinate x_1, x_2, ..., x_n. I suoi coefficienti a_{ij} sono i prodotti scalari $Ae_i \cdot e_j$, essendo e_1, e_2, ..., e_n la base ortonormale scelta.

È facile vedere che, per la simmetria della matrice A, vale per vettori arbitrari X e Y l'uguaglianza

$$AX \cdot Y = X \cdot AY.$$

Mostriamo anzitutto che la trasformazione A possiede almeno un autovalore reale e un autovettore corrispondente.

Consideriamo a tale scopo i valori della forma $AX \cdot X$ quando X percorre la sfera unitaria, cioè l'insieme dei vettori di lunghezza uno. La forma vi assume un massimo, che indichiamo con λ_1. Dimostriamo che questo massimo è un autovalore di A e che un vettore X_0 in cui esso è raggiunto è un autovettore corrispondente: $AX_0 = \lambda_1 X_0$.

Procederemo indirettamente, dimostrando che AX_0 è ortogonale a tutti i vettori ortogonali a X_0.

Per ogni vettore Z vale la disuguaglianza $AZ \cdot Z \leqslant \lambda_1 \mid Z \mid^2$. Per Z non nullo, infatti, $X = Z/|Z|$ è un vettore unitario e λ_1 è il massimo della forma $AX \cdot X$ sulla sfera unitaria; per Z nullo la disuguaglianza è immediata. Consideriamo $Z = X_0 + \epsilon Y$, dove ϵ è un numero reale e Y un vettore arbitrario ortogonale a X_0. Allora

$$AZ \cdot Z = (AX_0 + \epsilon AY) \cdot (X_0 + \epsilon Y) = AX_0 \cdot X_0 + 2\epsilon AX_0 \cdot Y + \epsilon^2 AY \cdot Y$$
$$= \lambda_1 + 2\epsilon AX_0 \cdot Y + \epsilon^2 AY \cdot Y.$$

Inoltre,

$$\mid Z \mid^2 = (X_0 + \epsilon Y) \cdot (X_0 + \epsilon Y) = \mid X_0 \mid^2 + \epsilon^2 \mid Y \mid^2 = 1 + \epsilon^2 \mid Y \mid^2,$$

poiché

$$X_0 \cdot Y = 0, \quad \mid X_0 \mid^2 = 1.$$

Pertanto,

$$\lambda_1 + 2\epsilon AX_0 \cdot Y + \epsilon^2 AY \cdot Y \leqslant \lambda_1 + \epsilon^2 \lambda_1 \mid Y \mid^2,$$

e dividendo per ϵ^2 otteniamo

$$(2/\varepsilon)\, AX_0 \cdot Y \leq \lambda_1 |Y|^2 - AY \cdot Y. \quad (14)$$

Quest'ultima disuguaglianza deve valere per ogni ε reale di valore assoluto sufficientemente piccolo.

Ciò è possibile soltanto se $AX_0 \cdot Y = 0$. Infatti, se $AX_0 \cdot Y > 0$, la disuguaglianza (14) è impossibile per ϵ positivo sufficientemente piccolo; se invece $AX_0 \cdot Y < 0$, essa è impossibile per ϵ negativo di valore assoluto sufficientemente piccolo. Dunque $AX_0 \cdot Y = 0$: AX_0 è effettivamente ortogonale a ogni vettore ortogonale a X_0. Ne segue che AX_0 e X_0 sono collineari, ossia $AX_0 = \lambda' X_0$, con λ′ reale. È facile verificare che $\lambda' = \lambda_1$, poiché

$$\lambda_1 = \boldsymbol{AX_0} \cdot \boldsymbol{X_0} = \lambda' \boldsymbol{X_0} \cdot \boldsymbol{X_0} = \lambda'.$$

È ora facile dimostrare che ogni forma quadratica reale può essere ridotta alla forma canonica mediante una trasformazione ortogonale.

Siano $e_1, e_2, \ldots, e_n$ la base ortonormale iniziale e $f_1, f_2, \ldots, f_n$ una nuova base ortonormale il cui primo vettore f_1 è l'autovettore X_0 di A. Indichiamo con $x_1, \ldots, x_n$ le coordinate di X nella base iniziale e con $x'_1, \ldots, x'_n$ quelle nella nuova base. Allora

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = P \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ \vdots \\ x'_n \end{bmatrix}$$

dove P è una matrice ortogonale.

Eseguiamo il cambiamento di variabili nella forma $AX \cdot X$. Nelle nuove variabili i suoi coefficienti sono $a'_{ij} = \boldsymbol{Af_i} \cdot \boldsymbol{f_j}$. Pertanto

$$a'_{11} = Af_1 \cdot f_1 = \lambda_1 f_1 \cdot f_1 = \lambda_1,$$
$$a'_{1j} = a'_{j1} = Af_1 \cdot f_j = \lambda_1 f_1 \cdot f_j = 0 \quad \text{per } j \neq 1.$$

cioè la forma diventa

$$\lambda_1 x'^2_1 + \phi(x'_2, \cdots, x'_n).$$

Abbiamo così separato, mediante una trasformazione ortogonale, il quadrato di una nuova variabile.

Ripetendo lo stesso ragionamento per la nuova forma $\phi(x'_2, \cdots, x'_n)$, e così via, otteniamo infine la forma canonica mediante una successione di trasformazioni ortogonali. Ma una successione di trasformazioni ortogonali equivale a un'unica trasformazione ortogonale. Il teorema è dimostrato.

§6. Funzioni di matrici e alcune loro applicazioni

Funzioni di matrici.

Le applicazioni dell'algebra lineare agli altri rami della matematica sono numerosissime e svariate. Non è esagerato affermare che le sue idee e i suoi risultati intervengono in gran parte della matematica contemporanea e della fisica teorica, in una forma o nell'altra, soprattutto attraverso il calcolo matriciale.

Esamineremo brevemente un metodo di applicazione del calcolo matriciale alla teoria delle equazioni differenziali ordinarie. Vi svolgono un ruolo importante le funzioni di matrici.

Definiamo anzitutto le potenze di una matrice quadrata A. Poniamo $A^0 = E$, $A^1 = A$, $A^2 = AA$, $A^3 = A^2A$, $A^4 = A^3A$, ecc. Dalla proprietà associativa segue facilmente che $A^mA^n = A^{m+n}$ per interi non negativi arbitrari m e n. Per le matrici abbiamo già definito la somma e il prodotto per un numero. Possiamo quindi attribuire in modo naturale un significato a un polinomio di una matrice: se $\phi(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$, poniamo per definizione $\phi(A) = a_0A^n + a_1A^{n-1} + \cdots + a_nE$. È così definita la più semplice funzione di argomento matriciale, il polinomio.

Mediante un passaggio al limite si può estendere facilmente il concetto di funzione di argomento matriciale a una classe assai più vasta dei polinomi in una variabile. Senza affrontare il problema in tutta la sua generalità, ci limiteremo qui alle funzioni analitiche.

Introduciamo anzitutto il limite di una successione di matrici. Si dice che la successione

$$A_1 = \begin{bmatrix} a_{11}^{(1)} & \cdots & a_{1n}^{(1)} \\ \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix}, \quad A_2 = \begin{bmatrix} a_{11}^{(2)} & \cdots & a_{1n}^{(2)} \\ \cdots & \cdots & \cdots \\ a_{n1}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix}, \cdots$$

converge alla matrice

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

(o ha per limite la matrice A) se $\lim_{k\to\infty} a_{ij}^{(k)} = a_{ij}$ per ogni i, j. La somma di una serie $A_1 + A_2 + \ldots + A_k + \ldots$ è definita come il limite delle somme parziali, $\lim_{k\to\infty} (A_1 + A_2 + \ldots + A_k)$, quando tale limite esiste.

Sia $f(z)$ una funzione analitica regolare in un intorno di $z = 0$. Essa si può sviluppare in serie di potenze:

$$f(z) = a_0 + a_1z + a_2z^2 + \cdots + a_kz^k + \cdots.$$

Per una matrice quadrata A è naturale porre

$$f(A) = a_0E + a_1A + a_2A^2 + \cdots + a_kA^k + \cdots.$$

Questa serie converge per tutte le matrici A i cui autovalori appartengono all'interno del cerchio di convergenza della serie di potenze $a_0 + a_1z + \ldots + a_kz^k + \ldots$.

Nelle applicazioni sono particolarmente interessanti le funzioni elementari di matrici.

Per esempio, la serie geometrica $E + A + A^2 + \ldots + A^k + \ldots$ converge quando tutti gli autovalori di A hanno modulo minore di 1; la sua somma è la matrice $(E - A)^{-1}$, in pieno accordo con la formula

$$1 + x + \cdots + x^k + \cdots = \frac{1}{1 - x}.$$

La rappresentazione di $(E - A)^{-1}$ mediante una serie infinita offre un metodo efficace per risolvere approssimativamente sistemi lineari la cui matrice dei coefficienti è vicina alla matrice identità.

Scrivendo infatti un tale sistema nella forma

$$(E - A)\,X = B,$$

otteniamo

$$X = (E - A)^{-1}\,B = B + AB + A^2B + \cdots \qquad (15)$$

Questa è una formula comoda per risolvere il sistema, purché la serie (15) converga abbastanza rapidamente.

È utile considerare anche la serie binomiale

$$(E + A)^m = E + \frac{m}{1}A + \frac{m(m-1)}{2!}A^2 + \cdots$$

che si può applicare, purché gli autovalori di A abbiano modulo minore di 1, non soltanto a esponenti naturali m, ma anche a esponenti frazionari e negativi.

Particolarmente importante nelle applicazioni è la funzione esponenziale di matrice:

$$e^A = E + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots.$$

La serie che definisce l'esponenziale converge per ogni matrice quadrata A. Questa funzione possiede proprietà analoghe a quelle dell'esponenziale ordinario. Per esempio, se A e B commutano, cioè $AB = BA$, allora $e^{A+B} = e^A \cdot e^B$. Per matrici che non commutano, invece, questa formula non vale in generale.

Applicazione ai sistemi di equazioni differenziali lineari ordinarie.

Nella teoria di tali sistemi è opportuno considerare matrici i cui elementi dipendono da una variabile indipendente:

$$U(t) = \begin{bmatrix} a_{11}(t) \cdots a_{1n}(t) \\ \cdots\cdots\cdots\cdots\cdots \\ a_{m1}(t) \cdots a_{mn}(t) \end{bmatrix}.$$

Per queste matrici si definisce naturalmente la derivata rispetto a t:

$$\frac{dU(t)}{dt} = \begin{bmatrix} a'_{11}(t) \cdots a'_{1n}(t) \\ \cdots\cdots\cdots\cdots\cdots \\ a'_{m1}(t) \cdots a'_{mn}(t) \end{bmatrix}.$$

Non è difficile verificare che alcune formule elementari di derivazione valgono anche per le matrici. Per esempio,

$$\frac{d(U+V)}{dt} = \frac{dU}{dt} + \frac{dV}{dt},$$

$$\frac{d(cU)}{dt} = c\,\frac{dU}{dt},$$

$$\frac{d(UV)}{dt} = \frac{dU}{dt}\,V + U\,\frac{dV}{dt}.$$

L'ordine dei fattori deve essere rigorosamente quello indicato nella formula!

Un sistema di equazioni differenziali lineari omogenee ordinarie

$$\begin{aligned} \frac{dy_1}{dt} &= a_{11}(t)\,y_1 + a_{12}(t)\,y_2 + \cdots + a_{1n}(t)\,y_n, \\ \frac{dy_2}{dt} &= a_{21}(t)\,y_1 + a_{22}(t)\,y_2 + \cdots + a_{2n}(t)\,y_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{dy_n}{dt} &= a_{n1}(t)\,y_1 + a_{n2}(t)\,y_2 + \cdots + a_{nn}(t)\,y_n \end{aligned}$$

si può scrivere, con questa notazione, nella forma

$$\frac{dY}{dt} = A(t)\,Y,$$

dove

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \qquad A(t) = \begin{bmatrix} a_{11}(t) \cdots a_{1n}(t) \\ \cdots\cdots\cdots\cdots \\ a_{n1}(t) \cdots a_{nn}(t) \end{bmatrix},$$

Si ottiene così una scrittura analoga a quella di una singola equazione differenziale lineare omogenea.

Se i coefficienti del sistema sono costanti, cioè se la matrice A è costante, anche la soluzione assomiglia formalmente a quella dell'equazione $y' = ay$. In questo caso, infatti, $Y = e^{At}C$, dove C è una colonna di costanti arbitrarie.

Questa forma della soluzione è molto comoda per i calcoli. Per ogni funzione analitica $f(z)$, definita sulle matrici considerate, vale infatti

$$f(B^{-1}LB) = B^{-1}f(L)\,B.$$

Poiché ogni matrice complessa può essere ridotta alla forma canonica di Jordan (si veda il §4), calcolare una funzione di una matrice arbitraria si riduce a calcolare una funzione di una matrice canonica, operazione agevole. Quindi, se $A = B^{-1}LB$, con L canonica, allora

$$Y = e^{At}C = B^{-1}e^{Lt}BC = B^{-1}e^{Lt}C',$$

dove $C' = BC$ è una colonna di costanti arbitrarie.

Da questa formula si ricavano facilmente espressioni esplicite per tutte le componenti della colonna incognita Y.

Il matematico sovietico I. A. Lappo-Danilevskij sviluppò con successo gli strumenti della teoria delle funzioni di matrici e fu il primo ad applicarli allo studio dei sistemi differenziali, anche con coefficienti variabili. L'autore annovera questi risultati fra le più brillanti conquiste matematiche dei cinquant'anni precedenti la stesura del libro.

Letture consigliate

G. Birkhoff and S. MacLane, A survey of modern algebra, Macmillan, New York, 1941.

V. N. Faddeeva, *Computational methods of linear algebra*. Traduzione di C. D. Benster, Dover, New York, 1959.

F. R. Gantmacher, *Applications of the theory of matrices*. Traduzione di J. L. Brenner, Interscience, New York, 1959.

--, *The theory of matrices*, voll. 1 e 2. Traduzione di K. A. Hirsch, Chelsea, New York, 1959.

[Questo libro e il precedente sono traduzioni dell'originale russo, pubblicato nel 1953, con varie aggiunte, alcune delle quali fornite dall'autore. I due volumi contengono molto materiale non disponibile negli altri testi dell'epoca.]

I. M. Gel'fand, Lectures on linear algebra, Interscience, New York, 1961.

F. Klein, *Elementary mathematics from an advanced standpoint*, vol. I: *Arithmetic, algebra, analysis*. Traduzione di E. R. Hedrick e C. A. Noble, Dover, New York, 1953.

O. Schreier and E. Sperner, Introduction to modern algebra and matrix theory, Chelsea, New York, 1951.

V. I. Smirnov, *Linear algebra and group theory*. Traduzione e revisione di R. A. Silverman, McGraw-Hill, New York, 1961.

H. W. Turnbull and A. C. Aitken, An introduction to the theory of canonical matrices, Dover, New York, 1961.

B. L. van der Waerden, *Modern algebra*, vol. I. Traduzione dalla seconda edizione tedesca riveduta di Fred Blum, con revisioni e aggiunte dell'autore. Frederick Ungar, New York, 1949.

Nota al capitolo XVI

[1] Il nome «equazione secolare» è nato nella meccanica celeste, in relazione al problema delle cosiddette perturbazioni secolari del moto dei pianeti.

Nota editoriale

Sono conservate le sei sezioni, le formule, i due schemi per il calcolo dei determinanti, la bibliografia e la nota. Le scritte inglesi interne alle formule sono tradotte; alcune espressioni sono ricomposte per correggere refusi o renderle leggibili. La barra sopra una matrice indica la trasposta; l'apice usato nella definizione del prodotto scalare complesso indica invece il coniugato.

Si mantiene la terminologia storica dell'autore: all'inizio «funzione lineare» comprende le funzioni oggi dette affini; «lineare omogenea» corrisponde all'uso moderno di «lineare». Gli «spazi affini» menzionati nel §2 sono qui spazi vettoriali privi di prodotto scalare, non gli spazi affini di punti nella terminologia moderna. Sono rese esplicite alcune ipotesi: sottospazi non vuoti, vettori non nulli nella definizione di angolo, basi ortonormali nelle trasformazioni ortogonali, regolarità locale per il teorema della funzione inversa e uso del campo complesso per la forma di Jordan. Per i prodotti scalari integrali si specifica una classe di funzioni sulla quale la positività sia effettiva.

Sono corretti indici e segni nelle coordinate, nello sviluppo del prodotto scalare, nella disuguaglianza di Schwarz, nel completamento dei quadrati e nella dimostrazione della riduzione ortogonale. Nella definizione di polinomio di matrice l'esponente del secondo termine è n − 1; nella definizione di limite viene reinserita l'uguaglianza al coefficiente limite; nella soluzione del sistema differenziale l'esponente è At. Le rotazioni piane senza autovettori reali hanno angolo non multiplo di 180°. La diagonalizzazione mediante autovettori è garantita dall'ipotesi di n autovalori distinti, mentre la forma di Jordan non richiede tale ipotesi.

CAPITOLO XVII

Geometria non euclidea

A. D. Aleksandrov

Da quando N. I. Lobačevskij dimostrò per la prima volta la possibilità di una geometria non euclidea e propose una nuova concezione del rapporto fra geometria e realtà materiale, il campo della geometria, i suoi metodi e le sue applicazioni si sono enormemente ampliati. Oggi i matematici studiano diversi «spazi»: oltre allo spazio euclideo, considerano lo spazio di Lobačevskij, quello proiettivo, vari spazi a *n* dimensioni e persino a infinite dimensioni, spazi riemanniani, topologici e altri ancora. Il loro numero è illimitato e ciascuno possiede proprietà proprie, una propria «geometria». In fisica si usano i cosiddetti «spazi» delle fasi e delle configurazioni; la teoria della relatività impiega il concetto di spazio curvo e altri risultati delle teorie geometriche astratte.

Come e da dove sono nate queste astrazioni matematiche? Quale fondamento reale, quale valore e quali applicazioni hanno? Qual è il loro rapporto con la realtà? Come vengono definite e studiate in matematica? Quale significato hanno, per la matematica, le idee generali della geometria contemporanea?

Risponderemo a queste domande nel presente capitolo. Non esporremo la teoria degli spazi matematici astratti in quanto tale: ciò richiederebbe una trattazione molto più lunga e un'attenzione assai maggiore ai suoi strumenti specifici. Vogliamo chiarire l'essenza delle nuove idee geometriche, ossia rispondere alle domande appena formulate; possiamo farlo senza dimostrazioni e formule complicate.

La storia del problema risale agli *Elementi* di Euclide e all'assioma, o postulato, delle parallele.

§1. Storia del postulato di Euclide

Negli *Elementi*, Euclide formula le premesse fondamentali della geometria sotto forma di postulati e assiomi. Fra essi figura il quinto postulato, indicato in altre versioni degli *Elementi* come undicesimo assioma. Nella formulazione oggi consueta esso afferma: «Per un punto esterno a una retta data si può condurre al più una retta parallela alla retta data». Ricordiamo che due rette sono parallele se giacciono in uno stesso piano e non si incontrano: intendiamo qui le rette intere, illimitate, non i loro segmenti finiti.

È facile dimostrare che per un punto A esterno a una retta a si può sempre condurre almeno una parallela ad a.

Abbassiamo infatti da A la perpendicolare b ad a e conduciamo per A la retta c perpendicolare a b (figura 1). La figura ottenuta è simmetrica rispetto a b, poiché gli angoli formati da b con a e c sono tutti retti. Ribaltando il piano attorno a b, le semirette di a e di c si sovrappongono alle rispettive semirette opposte. Se dunque a e c si incontrassero da una parte di AB, dovrebbero incontrarsi anche dall'altra. Avrebbero così due punti in comune, il che è impossibile: per una proprietà fondamentale della retta, per due punti passa una sola retta, perciò due rette con due punti comuni devono coincidere.

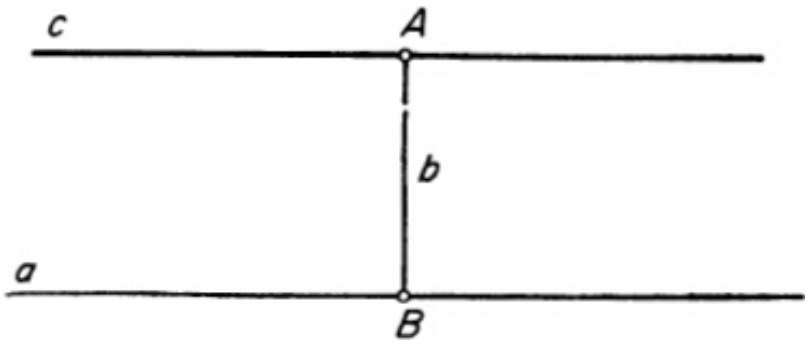


Figura 1.

Dalle proprietà fondamentali della retta e del movimento delle figure segue dunque che per un punto dato si può sempre condurre almeno una parallela a una retta data. Il ribaltamento attorno ad AB è infatti una rotazione del semipiano attorno a questa retta. Il postulato di Euclide completa il risultato affermando che tale parallela è unica e che non può essercene un'altra.

Fra gli altri postulati della geometria, questo occupa un posto particolare. La formulazione originale di Euclide è piuttosto complessa, ma anche quella consueta appena riportata presenta una difficoltà, già insita nel concetto di parallele: si considera la retta nella sua interezza. Come convincersi che due rette date siano parallele? Bisognerebbe prolungarle da entrambe le parti «all'infinito» e accertarsi che non si incontrino in nessun punto della loro estensione illimitata. La difficoltà è evidente. Probabilmente per questo il postulato delle parallele aveva una posizione particolare già per Euclide: negli *Elementi* viene usato soltanto a partire dalla proposizione 29, mentre le prime 28 ne fanno a meno. Data la sua complessità, era naturale desiderare di evitarlo. Già nell'antichità si tentò di cambiare la definizione delle

parallele o la formulazione del postulato, e soprattutto di dedurlo come teorema dagli altri assiomi e concetti fondamentali della geometria.

La teoria delle parallele, fondata sul quinto postulato, divenne così fin dall'antichità oggetto di commenti e critiche da parte di molti geometri. Lo scopo principale di queste ricerche era eliminare il quinto postulato, deducendolo come teorema dalle altre proposizioni fondamentali della geometria.

Il problema attirò molti studiosi: il greco Proclo, autore nel V secolo d.C. di un commento a Euclide; il persiano Nasir ad-Din at-Tusi[1], nel XIII secolo; l'inglese Wallis (1616-1703); l'italiano Saccheri (1667-1733); il filosofo e matematico tedesco Lambert (1728-1777); il francese Legendre (1752-1833), e molti altri. Nei più di duemila anni trascorsi dalla comparsa degli *Elementi*, essi gareggiarono in sottigliezza e ingegnosità geometrica nel tentativo di dimostrare il quinto postulato.

Il risultato, però, fu sempre negativo. Ogni volta si scopriva che l'autore della dimostrazione si era appoggiato a una proposizione, per quanto evidente, che non discendeva con necessità logica dalle altre premesse della geometria. In sostanza, aveva sostituito il quinto postulato con un'altra affermazione dalla quale esso seguiva, ma che richiedeva a sua volta una dimostrazione.[2]

Saccheri e Lambert penetrarono nel problema più profondamente degli altri. Saccheri tentò per primo una dimostrazione per assurdo: assunse l'affermazione contraria al quinto postulato e ne sviluppò le conseguenze, sperando di ottenere una contraddizione. Quando arrivò a risultati che gli sembravano del tutto inconcepibili, credette di avere risolto il problema. Si sbagliava: il contrasto con l'intuizione non costituisce ancora una contraddizione logica. Si trattava di dimostrare il postulato di Euclide sulla base delle altre proposizioni geometriche, non di convincersi ancora una volta della sua verità intuitiva. Il postulato è già di per sé intuitivamente persuasivo. Ma persuasività intuitiva e necessità logica sono cose diverse.

Lambert mostrò maggiore profondità di Saccheri e dei suoi predecessori. Seguendo la stessa strada, non trovò contraddizioni logiche e non commise i loro errori: non pretese di avere dimostrato il quinto postulato. Dopo di lui, tuttavia, all'inizio dell'Ottocento, Legendre tornò a «dimostrarlo» ricadendo nel vecchio errore: lo sostituì con altre affermazioni che richiedevano anch'esse una dimostrazione.

All'inizio del XIX secolo il problema rimaneva dunque irrisolto come ai tempi di Euclide. Gli sforzi erano stati vani e il problema sembrava resistere a ogni tentativo. Era un profondo enigma della geometria: per duemila anni non si era riusciti a risolvere una questione della cui risolubilità i migliori geometri non dubitavano.

La teoria delle parallele divenne uno dei problemi centrali della geometria ottocentesca. Vi si dedicarono Gauss, Lagrange, d'Alembert, Legendre, Wachter, Schweikart, Taurinus, Farkas Bolyai e altri.

La dimostrazione, però, non arrivava. La causa era l'insufficienza dei mezzi, oppure il problema era posto in modo inesatto? Questa domanda cominciò ad affacciarsi agli studiosi più profondi. Gauss, il celebre matematico tedesco, se ne occupava dal 1792 e comprese gradualmente come formulare correttamente la questione. Decise infine di abbandonare il quinto postulato e, dal 1813, sviluppò una serie di teoremi conseguenti all'affermazione opposta. Poco dopo seguirono la stessa strada Schweikart, allora professore di diritto a Charkiv, e Taurinus. Nessuno di loro diede però una risposta definitiva. Gauss tenne accuratamente nascoste le proprie ricerche; Schweikart si limitò alla corrispondenza privata con lui; soltanto Taurinus pubblicò elementi di una nuova geometria fondata sulla negazione del quinto postulato, pur escludendo egli stesso che tale geometria fosse possibile. Restava quindi senza risposta la domanda sulla correttezza del problema. A rispondere per primo fu N. I. Lobačevskij, giovane professore dell'Università di Kazan': il 23 febbraio 1826 presentò una relazione sulla teoria delle parallele alla facoltà fisico-matematica e nel 1829 ne pubblicò il contenuto nella rivista dell'Università.

§2. La soluzione di Lobačevskij

1. Le idee di Lobačevskij.

L'essenza della sua soluzione del problema del quinto postulato è espressa nei *Nuovi principi della geometria* (1835) con queste parole:

«È noto che nella geometria la teoria delle parallele è rimasta finora incompleta. Gli inutili sforzi compiuti per duemila anni, a partire da Euclide, mi hanno indotto a sospettare che i concetti stessi non contengano la verità che si è voluta dimostrare, ma che essa possa essere verificata soltanto mediante esperimenti, come le altre leggi fisiche, per esempio attraverso osservazioni astronomiche. Convinto infine della verità della mia congettura e considerando completamente risolto il difficile problema, nel 1826 misi per iscritto i miei ragionamenti».

Esaminiamo il senso di questa affermazione, nella quale si concentra la nuova idea di Lobačevskij. Essa non risolse soltanto il problema del quinto postulato, ma rivoluzionò tutta la nostra concezione della geometria e anche di altri rami della matematica.

Lobačevskij aveva cominciato a occuparsi della teoria delle parallele già nel 1815, tentando dapprima, come gli altri geometri, di dimostrare il quinto postulato. Nel 1823 comprese chiaramente che tutte quelle dimostrazioni, «di qualunque genere siano, possono essere considerate soltanto chiarimenti, ma non meritano il nome di dimostrazioni matematiche nel pieno senso della parola».[3] Aveva dunque riconosciuto che «i concetti stessi non contengono la verità che si è voluta dimostrare»: il quinto postulato non è deducibile dalle proposizioni e dai concetti fondamentali della geometria. Come si convinse di tale impossibilità?

Proseguì sulla strada di cui Saccheri e Lambert avevano percorso i primi passi. Assunse come ipotesi l'affermazione contraria al postulato euclideo: «Per un punto esterno a una retta data si possono condurre non una, ma almeno due rette parallele alla retta data». Accettiamola provvisoriamente come assioma, aggiungiamola alle altre proposizioni geometriche e sviluppiamone le conseguenze. Se essa è incompatibile con le altre premesse, arriveremo a una contraddizione e avremo dimostrato indirettamente il quinto postulato. Ma poiché nessuna contraddizione si manifestava, si delinearono le due conclusioni alle quali giunse Lobačevskij.

La prima era che il quinto postulato non è dimostrabile. La seconda era che, sulla base dell'assioma opposto appena formulato, si può sviluppare una catena di conseguenze, cioè di teoremi, priva di contraddizioni. Queste conseguenze costituiscono una teoria logicamente possibile e coerente, interpretabile come una nuova geometria non euclidea. Lobačevskij la chiamò prudentemente «immaginaria», perché non ne aveva ancora trovato un'interpretazione reale. Ma era convinto della sua possibilità logica. Nell'esprimere e difendere questa convinzione mostrò la grandezza di chi sostiene le proprie idee senza esitazioni e non le nasconde per timore delle incomprensioni e delle critiche.

Le prime due conclusioni di Lobačevskij furono dunque l'indimostrabilità del quinto postulato e la possibilità di costruire, sull'assioma contrario, una geometria logicamente ricca e compiuta quanto quella euclidea, pur in contrasto con l'immagine intuitiva dello spazio. Egli sviluppò effettivamente questa geometria, che oggi porta il suo nome. Il risultato generale era di enorme importanza: è logicamente concepibile più di una geometria. Ne discuteremo più avanti l'intero significato; esso contiene già una parte sostanziale della risposta alle domande sugli spazi astratti poste all'inizio del capitolo.

Torniamo però all'affermazione citata. Lobačevskij dice che la verità geometrica, come quella delle altre leggi fisiche, si può verificare soltanto sperimentalmente. La verità va quindi intesa come corrispondenza fra concetti astratti e realtà. Tale corrispondenza si stabilisce mediante l'esperienza: per verificare un risultato occorrono indagini sperimentali, e la sola deduzione logica non basta. La geometria euclidea riflette con grande precisione le proprietà reali dello spazio, ma non si può escludere che ricerche ulteriori mostrino che essa è soltanto approssimativamente vera come teoria dello spazio reale. In tal caso la geometria, intesa come scienza dello spazio reale e non come sistema logico, dovrebbe essere modificata e precisata in accordo con i nuovi dati sperimentali.

Questa brillante idea di Lobačevskij ha trovato piena conferma in un nuovo ramo della fisica: la teoria della relatività.

Lobačevskij stesso eseguì calcoli basati su osservazioni astronomiche per verificare l'esattezza della geometria euclidea. Essi la confermarono entro i limiti di precisione allora disponibili. La situazione è poi mutata, ma va subito sottolineato che neppure la geometria di Lobačevskij si è rivelata una descrizione più esatta dello spazio fisico, le cui proprietà sono risultate diverse e più complesse. Già prima, tuttavia, essa aveva trovato un fondamento e applicazioni in un altro ambito, di cui parleremo più avanti.

Lobačevskij non considerava affatto la propria geometria un semplice schema logico costruito su premesse arbitrarie. Il compito principale, per lui, era studiare il rapporto fra geometria e realtà, non soltanto analizzarne logicamente i fondamenti. Poiché nessun esperimento può decidere con precisione assoluta la verità del postulato di Euclide, ha senso esplorare le possibilità logiche contenute nelle premesse geometriche fondamentali. Questa ricerca matematica aiuta a indicare le vie lungo le quali può procedere lo studio fisico dello spazio. Inoltre, la geometria euclidea è un caso limite di quella di Lobačevskij, che comprende quindi possibilità più ampie. Limitarsi al postulato euclideo avrebbe ostacolato lo sviluppo della teoria. La teoria deve oltrepassare i confini del noto per cercare nuove leggi e nuovi fatti. Comprendere più profondamente i legami fra matematica e realtà consente di scegliere, fra le diverse possibilità logiche, quelle che hanno maggiori probabilità di essere utili nello studio della natura. Se dopo Lobačevskij la geometria non avesse sviluppato la teoria matematica delle possibili proprietà dello spazio, la fisica contemporanea non avrebbe avuto gli strumenti necessari per formulare e sviluppare la relatività.

Possiamo dunque riassumere così la soluzione di Lobačevskij:

1. Il postulato non è dimostrabile.

2. Aggiungendo l'assioma contrario alle altre proposizioni fondamentali si può sviluppare una geometria completa e logicamente coerente, diversa da quella euclidea.

3. La verità dei risultati di una geometria logicamente concepibile, quando la si applica allo spazio reale, può essere verificata soltanto sperimentalmente. Una geometria non va sviluppata soltanto come schema logico arbitrario, ma anche come teoria capace di indicare possibili vie e metodi per elaborare le teorie fisiche.

Questa soluzione era del tutto diversa da quella cercata dai geometri che volevano dimostrare il postulato euclideo. Contrastava tanto con le idee consolidate da incontrare scarsa comprensione fra i matematici: era troppo nuova e radicale. Lobačevskij, per così dire, tagliò il nodo gordiano della teoria delle parallele, anziché cercare di scioglierlo come gli altri si aspettavano.

2. Altri geometri e filosofi.

Quasi contemporaneamente, il geometra ungherese János Bolyai (1802-1860) scoprì anch'egli l'impossibilità di dimostrare il quinto postulato e la possibilità di una geometria non euclidea. Pubblicò i risultati nel 1832, in appendice al trattato geometrico del padre Farkas. Questi aveva precedentemente inviato il lavoro del figlio a Gauss per averne un parere, ricevendo una risposta incoraggiante nella quale Gauss dichiarava di essere giunto da tempo agli stessi risultati. Gauss, tuttavia, non pubblicò nulla; in una lettera spiegò di temere di essere frainteso.

Nella scienza accade spesso che i tempi siano maturi per certi risultati e che più studiosi li raggiungano quasi simultaneamente e indipendentemente. Il calcolo differenziale e integrale fu sviluppato da Newton e Leibniz; Wallace giunse indipendentemente a idee analoghe a quelle di Darwin; elementi della teoria della relatività furono elaborati da Einstein e anche da Poincaré. Gli esempi sono molti. Essi mostrano ancora una volta che la scienza cresce risolvendo problemi ormai maturi, e non soltanto attraverso scoperte e intuizioni accidentali. Così avvenne anche per la scoperta della possibilità di una geometria non euclidea, raggiunta da più geometri.

Ciò non toglie che si possa valutare il contributo individuale. Conta la priorità, ma contano anche la chiarezza e la profondità dei risultati, la loro coerenza e la solidità degli argomenti. Né Schweikart né Taurinus erano convinti della pari dignità della nuova geometria: un punto decisivo, tanto più che risultati parziali erano già stati ottenuti da Saccheri e Lambert. Gauss, pur avendo apparentemente tale convinzione, non fu abbastanza risoluto da esporsi pubblicamente.

Bolyai non mostrò questa esitazione, ma non sviluppò le nuove idee con l'ampiezza e la profondità di Lobačevskij. Questi le espose apertamente per primo, oralmente nel 1826 e a stampa nel 1829, e continuò a svilupparle e diffonderle in numerosi lavori, fino alla *Pangeometria* del 1855. Ormai anziano e cieco, la dettò conservando intatte la forza del pensiero e la fiducia nella propria opera. Per questo la nuova geometria porta il suo nome.

§3. La geometria di Lobačevskij

1. Alcuni risultati sorprendenti.

Lobačevskij partì dunque dall'affermazione contraria al quinto postulato: in un piano, per un punto esterno a una retta si possono condurre almeno due rette che non la incontrano. Ne dedusse conseguenze profonde, che costituirono la nuova geometria. Essa venne costruita come teoria possibile, come insieme di teoremi deducibili logicamente da quel postulato insieme alle altre ipotesi fondamentali della geometria euclidea, o «abituale», come la chiamava Lobačevskij.[4]

Egli ottenne risultati analoghi a quelli della geometria elementare «abituale», fino alla trigonometria non euclidea, alla risoluzione dei triangoli e al calcolo di aree e volumi. Non possiamo seguire qui tutta la catena delle sue deduzioni, per ragioni di spazio, non perché sia troppo complicata. Anche un corso scolastico di geometria ordinaria è lungo, e le deduzioni di Lobačevskij non sono né più semplici né più brevi. Citeremo soltanto alcuni risultati sorprendenti, rimandando alla letteratura specifica chi desideri approfondire. Più avanti presenteremo una semplice interpretazione concreta della geometria non euclidea.

Cominciamo dalle parallele. Siano dati una retta a e un punto A esterno a essa; abbassiamo da A la perpendicolare AB ad a. Per l'ipotesi fondamentale esistono almeno due rette per A che non incontrano a. Neppure le rette comprese nell'angolo fra esse incontrano a. Nella figura 2, è vero, le rette b e b', prolungate abbastanza, incontrerebbero effettivamente a, in contrasto con l'ipotesi. Non c'è da stupirsi: Lobačevskij non ragionava sulle figure disegnate nel piano ordinario, ma sviluppava le conseguenze logiche di un'ipotesi contraria a ciò che siamo abituati a vedere nei disegni. Le figure hanno qui soltanto una funzione ausiliaria. Non possono esprimere fedelmente i fatti della geometria non euclidea, perché vi tracciamo rette ordinarie in un piano ordinario, euclideo entro la precisione del disegno.

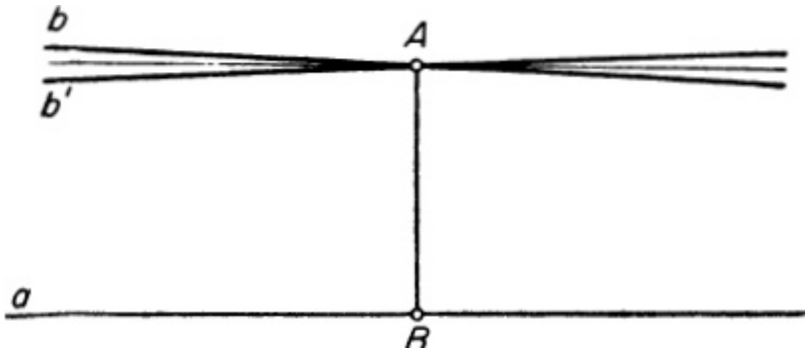


Figura 2.

Il contrasto fra possibilità logica e rappresentazione visiva fu un ostacolo importante alla comprensione della geometria di Lobačevskij. Ma quando si considera la geometria come teoria logica, si deve chiedere rigore logico ai ragionamenti, non accordo con i disegni consueti.

2. L'angolo di parallelismo.

Torniamo alla retta *a* e al punto *A*. Per *A* tracciamo una semiretta *x* che non incontri *a*, per esempio perpendicolare ad *AB*. Facciamola ruotare attorno ad *A*, diminuendo l'angolo ϕ fra *AB* e *x*, ma senza farle incontrare *a*. Essa raggiunge una posizione limite, corrispondente al minimo valore di ϕ. Anche questa semiretta limite *c* non incontra *a*.

Se infatti la incontrasse in un punto *X* (figura 3), potremmo scegliere un punto *X′* più a destra e ottenere una semiretta *AX′* che incontra *a* formando con *AB* un angolo maggiore. È impossibile, perché per la costruzione di *c* ogni semiretta che forma con *AB* un angolo maggiore non incontra *a*.

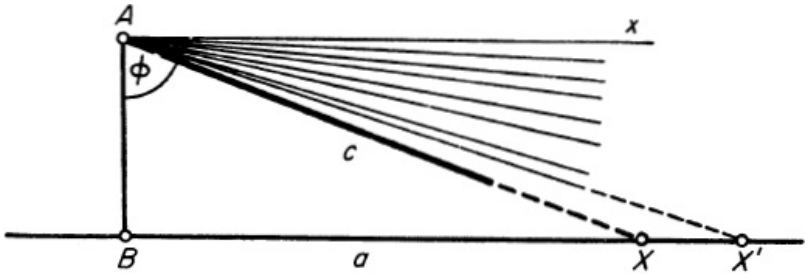


Figura 3.

Dunque *c* non incontra *a* ed è la posizione estrema delle semirette uscenti da *A* che non la incontrano, dal lato considerato.

Per simmetria esiste dall'altra parte una semiretta *c′* con la stessa proprietà. Se *c* e *c′* fossero l'una il prolungamento dell'altra, formerebbero un'unica retta *c* + *c′*. Essa sarebbe la sola parallela ad *a* per *A*: alla minima rotazione, una delle due semirette incontrerebbe *a*. Poiché abbiamo assunto che le parallele siano almeno due, *c* e *c′* non possono essere prolungamenti reciproci.

Abbiamo così dimostrato il primo teorema della geometria di Lobačevskij:

Per un punto *A* esterno a una retta *a* si possono condurre due semirette *c* e *c′* che non incontrano *a*, mentre ogni semiretta interna all'angolo fra esse rivolto verso *a* la incontra.

Prolungando *c* e *c′*, otteniamo due rette (figura 4) che non incontrano *a*. Ogni retta per *A* compresa nell'angolo α non incontra *a*, mentre ogni retta compresa nell'angolo β la incontra. Lobačevskij chiamò *c* e *c′* parallele ad *a*: la prima verso destra, la seconda verso sinistra. La metà dell'angolo β si chiama angolo di parallelismo; è minore di un angolo retto, poiché β è minore di due retti.

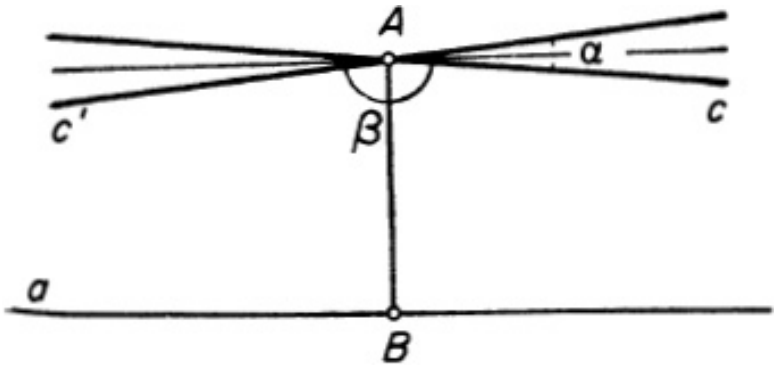


Figura 4.

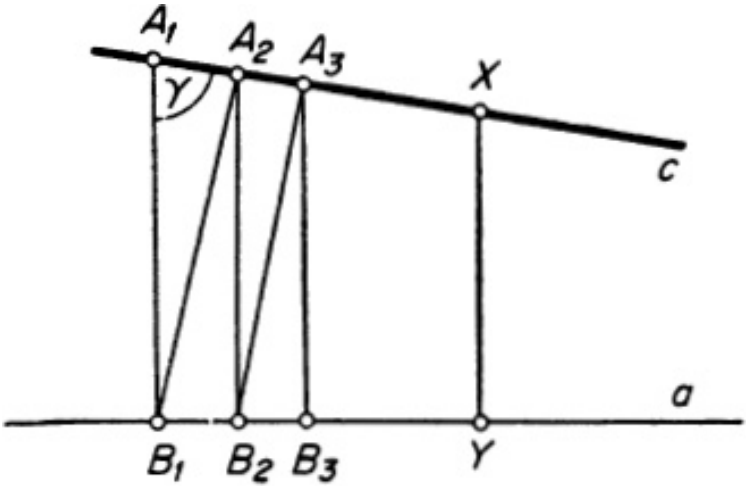


Figura 5.

3. Avvicinamento delle parallele; la curva equidistante.

Studiamo ora come varia la distanza da *a* di un punto *X* che si muove su *c* (figura 5). Nella geometria euclidea la distanza fra parallele è costante. Qui, invece, spostando *X* verso destra, la sua distanza da *a*, cioè la lunghezza della perpendicolare *XY*, diminuisce.

Da un punto A_1 abbassiamo la perpendicolare A_1B_1 ad *a*; da B_1 abbassiamo la perpendicolare B_1A_2 a *c*. Poiché γ è acuto, A_2 si trova a destra di A_1. Da A_2 abbassiamo infine la perpendicolare A_2B_2 ad *a*. Mostriamo che essa è più corta di A_1B_1.

Il teorema secondo cui il segmento perpendicolare è più corto di ogni segmento obliquo condotto dallo stesso punto alla retta vale anche nella geometria di Lobačevskij: la dimostrazione scolastica non impiega le parallele né risultati che ne dipendano. Quindi B_1A_2, perpendicolare a *c*, è più corto di A_1B_1; analogamente A_2B_2, perpendicolare ad *a*, è più corto di B_1A_2. A maggior ragione, dunque, A_2B_2 è più corto di A_1B_1.

Abbassando da B_2 la perpendicolare B_2A_3 a *c* e ripetendo il ragionamento, troviamo che A_3B_3 è più corto di A_2B_2. Otteniamo così una successione di perpendicolari sempre più corte: le distanze di A_1, A_2, ... da *a* diminuiscono. Completando il ragionamento si dimostra in generale che, se X'' si trova su *c* a destra di X', allora $X''Y''$ è più corto di $X'Y'$. Non ci soffermiamo su

questo punto: il ragionamento precedente chiarisce il fenomeno e una dimostrazione rigorosa non rientra nei nostri scopi.

Si può dimostrare un fatto ancora più notevole: quando *X* si sposta a destra lungo *c*, la distanza *XY* non soltanto diminuisce, ma tende a zero quando *X* va all'infinito. Le parallele *a* e *c* si avvicinano dunque asintoticamente! Nella direzione opposta, invece, la loro distanza cresce senza limite.

Nella geometria euclidea una parallela resta a distanza costante dalla retta data. Nella geometria di Lobačevskij non esistono coppie di rette distinte con questa proprietà: ogni retta si allontana indefinitamente dall'altra da una parte o da entrambe. Il luogo dei punti a distanza costante da una retta, da un lato fissato, non è quindi una retta, ma una curva detta equidistante.

Queste conclusioni sono sorprendenti e del tutto incompatibili con la rappresentazione visiva abituale. Ma, come abbiamo già detto, tale discordanza non è un argomento contro una teoria astratta sviluppata logicamente dalle premesse assunte.

4. Grandezza dell'angolo di parallelismo.

Studiamo l'angolo γ che la parallela *c* alla retta *a* forma con la perpendicolare *CA* (figura 6). Mostriamo che esso diminuisce quando aumenta la distanza di *C* da *a*. Dimostriamo prima questo fatto: se due rette *b* e *b′* formano angoli alterni interni uguali α e α′ con una trasversale *BB′*, esse possiedono una perpendicolare comune (figura 7).

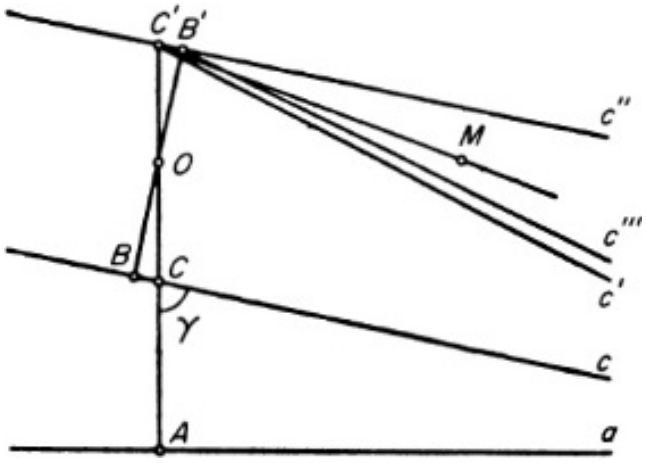


Figura 6.

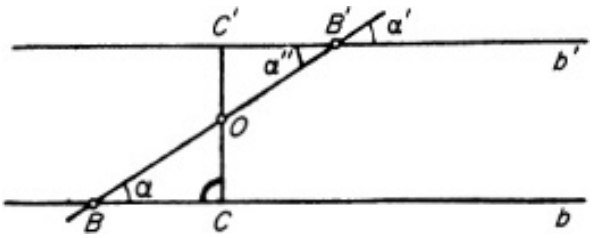


Figura 7.

Per il punto medio *O* di *BB′* conduciamo *CC′* perpendicolare a *b*. Nei triangoli *OBC* e *OB′C′*, i lati *OB* e *OB′* sono uguali per costruzione; gli angoli in *O* sono uguali perché opposti al vertice. Anche α″ e α′ sono opposti al vertice, quindi uguali; ma α′ = α per ipotesi, dunque α″ = α. I due triangoli hanno così un lato e i due angoli adiacenti rispettivamente uguali e sono congruenti. In particolare, sono uguali gli angoli in *C* e *C′*. Il primo è retto per costruzione, quindi lo è anche il secondo: *CC′* è perpendicolare sia a *b* sia a *b′*. L'esistenza della perpendicolare comune è dimostrata.

Proviamo ora che l'angolo di parallelismo diminuisce con la distanza dalla retta. Se *C′* è più lontano da *a* di *C*, la parallela *c′* per *C′* forma con *C′A* un angolo minore di quello formato da *c* per *C*, come nella figura 6.

Per *C′* tracciamo una retta *c″* che formi con *C′A* lo stesso angolo γ formato da *c*. Le rette *c* e *c″* formano angoli uguali con *CC′* e hanno quindi una perpendicolare comune *BB′*. Per *B′* possiamo condurre una parallela limite *c*□ a *c*, che forma con *BB′* un angolo minore di un retto. Scegliendo un punto *M* nella regione compresa fra *c″* e *c*□, otteniamo una semiretta *C′M* rivolta verso destra che non incontra *c*□, né quindi *c* e, a maggior ragione, *a*. Essa forma con *C′A* un angolo minore di γ. La parallela limite *c′* ad *a*, essendo la posizione estrema delle semirette per *C′* che non incontrano *a*, forma un angolo ancora minore. L'angolo di parallelismo in *C′* è dunque minore di quello in *C*, come volevamo dimostrare.

Abbiamo mostrato che l'angolo di parallelismo diminuisce al crescere della distanza di *C* da *a*. Si può dimostrare di più: quando *C* si allontana all'infinito, l'angolo tende a zero. A distanza abbastanza grande da *a*, la parallela forma quindi con la perpendicolare un angolo arbitrariamente piccolo.[5] In altre parole, se una perpendicolare ad *a* viene inclinata di un angolo piccolissimo attorno a un punto abbastanza lontano, la retta inclinata non incontra più *a*. Anche questo fatto è sorprendente, ma ne incontreremo altri non meno singolari.

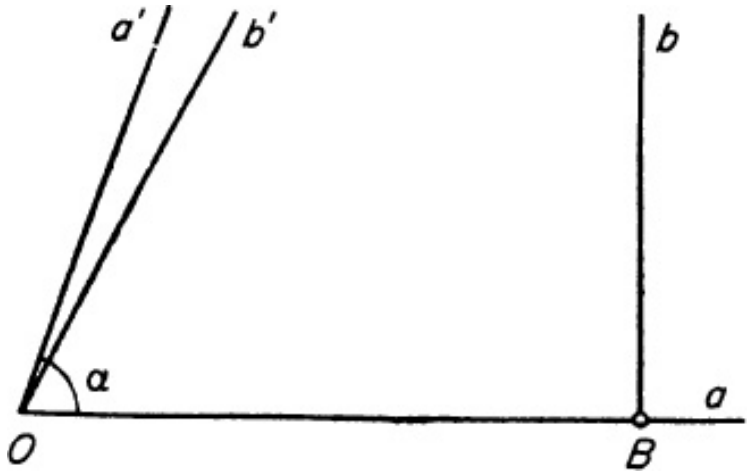


Figura 8.

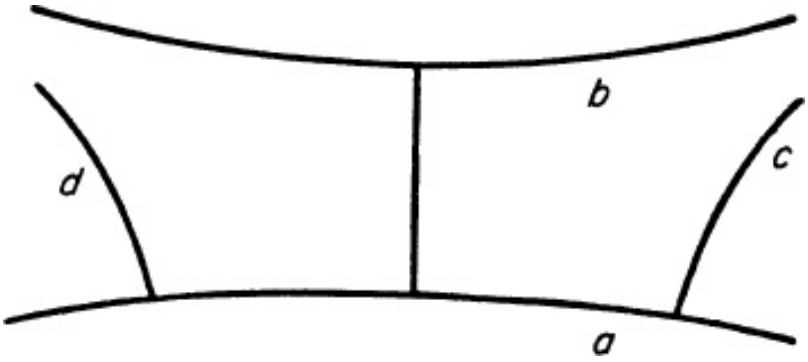


Figura 9.

Consideriamo, per esempio, un angolo acuto α formato dalle semirette *a* e *a′*. In un punto *B* di *a* abbastanza lontano dal vertice *O*, innalziamo la perpendicolare *b*, in modo che l'angolo di parallelismo corrispondente a *OB* sia minore di α (figura 8). La parallela *b′* a *b* condotta per *O* forma allora con *a* un angolo minore di α. Essa non incontra *b*, e a maggior ragione non la incontra *a′*. Dunque una perpendicolare a un lato di un angolo acuto, innalzata abbastanza lontano dal vertice, non incontra l'altro lato.

5. Altri teoremi sorprendenti.

Abbiamo sviluppato le deduzioni precedenti con un duplice scopo. Anzitutto volevamo mostrare, con esempi semplici, come i teoremi della geometria di Lobačevskij si ricavino effettivamente dalle premesse assunte. È un esempio elementare del modo di ragionare nella geometria astratta e della possibilità di ottenere conclusioni indipendenti dalla rappresentazione visiva abituale. In secondo luogo, volevamo mostrare la singolarità dei risultati di questa geometria. Eccone altri.

Due rette distinte del piano di Lobačevskij possono incontrarsi; oppure essere parallele nel senso di Lobačevskij, avvicinandosi asintoticamente da una parte e allontanandosi senza limite dall'altra; oppure possedere una perpendicolare comune e allontanarsi senza limite da entrambe le parti di essa.

Se *a* e *b* hanno una perpendicolare comune (figura 9), si possono condurre due perpendicolari *c* e *d* ad *a* che siano parallele, nel senso di Lobačevskij, a *b*. Tutta la retta *b* resta allora nella striscia fra *c* e *d*.

Il limite di circonferenze di raggio crescente indefinitamente, mantenute tangenti in un punto fissato, non è una retta, ma una curva detta circonferenza limite, o orociclo. Per tre punti non allineati non passa sempre una circonferenza; passa però una circonferenza, oppure un orociclo, oppure un'equidistante, cioè una curva i cui punti hanno distanza costante da una retta.

La somma degli angoli di un triangolo è sempre minore di due angoli retti. Se un triangolo si ingrandisce in modo che tutte e tre le altezze crescano senza limite, i suoi tre angoli tendono a zero.

Non esistono triangoli di area arbitrariamente grande.

Due triangoli con angoli rispettivamente uguali sono congruenti.

La lunghezza l di una circonferenza non è proporzionale al raggio r, ma cresce più rapidamente, essenzialmente secondo una legge esponenziale. Vale infatti la formula

$$l = \pi k\,(e^{r/k} - e^{-r/k}), \qquad (1)$$

dove k è una costante il cui valore numerico dipende dall'unità di lunghezza. Poiché

$$e^{r/k} = 1 + r/k + \tfrac{1}{2}(r/k)^2 + \ldots,$$
$$e^{-r/k} = 1 - r/k + \tfrac{1}{2}(r/k)^2 - \ldots$$

dalla (1) otteniamo:

$$l = 2\pi r\left(1 + \frac{1}{6}\frac{r^2}{k^2} + \cdots\right). \qquad (2)$$

Soltanto quando il rapporto r/k è piccolo vale con buona approssimazione $l = 2\pi r$.

Tutti questi risultati sono conseguenze logiche delle premesse assunte: l'assioma di Lobačevskij insieme alle altre proposizioni fondamentali della geometria «abituale».

6. Confronto fra le geometrie di Lobačevskij e di Euclide.

Una proprietà importantissima della geometria di Lobačevskij è che, in regioni abbastanza piccole, essa differisce poco da quella euclidea: quanto più piccola è la regione, tanto minore è la differenza. Per triangoli sufficientemente piccoli, per esempio, i rapporti fra lati e angoli sono descritti con buona precisione dalle formule della trigonometria ordinaria, tanto meglio quanto più piccolo è il triangolo.

La formula (2) mostra che, per piccoli raggi, la lunghezza della circonferenza è con buona approssimazione proporzionale al raggio. Analogamente, la somma degli angoli di un triangolo differisce poco da due angoli retti, e così via.

Nella formula della circonferenza compare una costante k, dipendente dall'unità di lunghezza. Se il raggio è piccolo rispetto a k, cioè se r/k è piccolo, la (2) mostra che l è vicina a $2\pi r$. In generale, quanto minore è il rapporto fra le dimensioni della figura e questa costante, tanto più le sue proprietà si avvicinano a quelle della figura corrispondente nella geometria euclidea.[6]

Il rapporto r/k fornisce dunque un parametro per lo scostamento dalle proprietà euclidee, quando r misura le dimensioni della figura: raggio della circonferenza, lati del triangolo e così via.

Ne segue una conseguenza importante.

Supponiamo di considerare lo spazio reale del mondo esterno, misurando le distanze in chilometri, e supponiamo che k sia molto grande, per esempio 10^{12}.

Anche per una circonferenza di raggio 100 km, la (2) darebbe allora un rapporto fra lunghezza e raggio diverso da 2π per meno di 10^{-9}. Anche le altre relazioni euclidee presenterebbero scostamenti piccolissimi. Su una scala di un chilometro il parametro r/k sarebbe 10^{-12}; su quella di un metro sarebbe 10^{-15}. Le correzioni relative nella formula della circonferenza sono ancora minori, dell'ordine di $(r/k)^2$. Si tratterebbe quindi di deviazioni trascurabili per le misure ordinarie. Per confronto, le dimensioni di un atomo sono dell'ordine di 10^{-13} km. Su scala astronomica, invece, il rapporto r/k potrebbe non essere così piccolo.

Per questo Lobačevskij suppose che, pur essendo la geometria euclidea molto esatta su scala ordinaria, le osservazioni astronomiche potessero rivelarne gli scostamenti. Come abbiamo già detto, l'idea generale è stata confermata; le piccole deviazioni osservate su scala astronomica risultano però più complesse di quelle descritte dalla sua geometria.

Il ragionamento conduce anche a un'altra conclusione importante. Poiché lo scostamento dalla geometria euclidea diminuisce al crescere di k, quando k tende all'infinito la geometria di Lobačevskij passa al limite in quella di Euclide. Quest'ultima è dunque un suo caso limite. Se si include questo caso, la teoria comprende anche la geometria euclidea ed è, in tal senso, più generale. Per questo Lobačevskij la chiamò «pangeometria», ossia geometria universale. Un simile rapporto fra teorie ricorre continuamente nella matematica e nelle scienze naturali: la nuova teoria include la precedente come caso limite, seguendo il progresso della conoscenza da risultati particolari a risultati più generali.

Tuttavia, tutti questi ragionamenti resterebbero un gioco del pensiero difficilmente comprensibile se non si potesse dare alla geometria di Lobačevskij un significato concreto relativamente semplice, entro il sistema dei concetti euclidei consueti. Lobačevskij non risolse definitivamente questo problema: vi riuscirono i suoi successori, quasi quarant'anni dopo la pubblicazione del suo primo lavoro. La soluzione è descritta nel paragrafo seguente.

§4. Il significato concreto della geometria di Lobačevskij

1. L'interpretazione di Beltrami sulla pseudosfera.

La prima interpretazione intuitiva fu data nel 1868 dal geometra italiano Beltrami. Egli osservò che la geometria intrinseca di una particolare superficie, la pseudosfera, coincide con quella di una parte del piano di Lobačevskij. Ricordiamo che per geometria intrinseca di una superficie si intende l'insieme delle proprietà delle figure che si possono determinare misurando lunghezze soltanto sulla superficie stessa. A sinistra nella figura 10 è disegnata la trattrice: una curva per la quale il segmento di tangente compreso fra il punto di tangenza e l'asse *Oy* ha lunghezza costante. L'asse *Oy* ne è un asintoto. Facendo ruotare la trattrice attorno a questo asintoto, si ottiene la pseudosfera, rappresentata a destra.

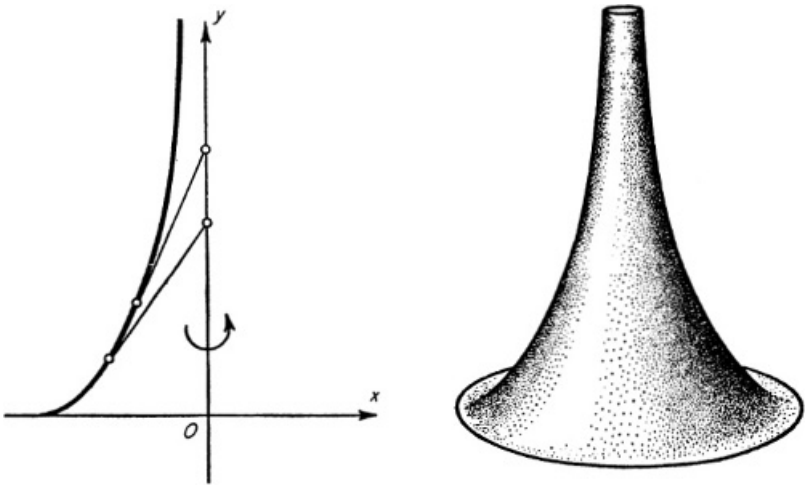


Figura 10.

L'interpretazione di Beltrami consiste nel far corrispondere tutte le relazioni geometriche di una parte del piano di Lobačevskij a quelle di una parte opportuna della pseudosfera, con le seguenti convenzioni. I segmenti rettilinei sono rappresentati dalle linee di minima lunghezza sulla superficie, dette geodetiche; la distanza fra due punti è la lunghezza del percorso più breve che li unisce sulla superficie. Due figure sono congruenti se si può stabilire una corrispondenza fra i loro punti che conservi tutte le distanze intrinseche. Un movimento sulla pseudosfera che conservi le dimensioni intrinseche della figura, pur deformandola per flessione nello spazio ambiente, rappresenta un movimento nel piano di Lobačevskij. Lunghezze, angoli e aree si misurano sulla superficie nel modo consueto e corrispondono alle rispettive grandezze della geometria non euclidea.

Con queste convenzioni, ogni proposizione della geometria di Lobačevskij relativa alla porzione di piano considerata corrisponde a un fatto della geometria intrinseca della pseudosfera. La geometria di Lobačevskij acquista così un significato concreto: è una descrizione astratta della geometria della pseudosfera, entro i limiti della porzione rappresentata.

Trent'anni prima della scoperta di Beltrami, F. Minding aveva già studiato la geometria intrinseca della pseudosfera, stabilendo proprio le proprietà che ne mostrano la coincidenza con quella di Lobačevskij. Né lui né altri riconobbero però questa connessione, finché le nuove idee non si furono diffuse abbastanza. A Beltrami bastò confrontare i risultati di Lobačevskij e di Minding per accorgersene.

La scoperta di Beltrami cambiò immediatamente l'atteggiamento dei matematici: da «fittizia», la geometria di Lobačevskij divenne concreta.[7]

2. L'interpretazione di Klein nel cerchio e nella sfera.

Come abbiamo sottolineato, sulla pseudosfera è rappresentata soltanto una parte del piano di Lobačevskij, non il piano intero.[8]

Restava quindi da interpretare concretamente la geometria dell'intero piano e, a maggior ragione, quella dello spazio. Vi riuscì in seguito, nel 1870, il matematico tedesco Klein. Vediamo in che cosa consiste la sua soluzione.

In un piano euclideo ordinario prendiamo un cerchio e consideriamone soltanto l'interno, escludendo la circonferenza e la regione esterna. Chiameremo «piano» questo interno, perché rappresenterà un piano di Lobačevskij. Chiameremo «rette» le corde, escluse le estremità sulla circonferenza. Infine, chiameremo «movimento» ogni trasformazione biunivoca del cerchio aperto in sé che porti rette in rette, cioè corde in corde. L'esempio più semplice è la rotazione attorno al centro, ma esistono molte altre trasformazioni di questo tipo, che descriveremo fra poco.

Con queste convenzioni terminologiche, i fatti della geometria ordinaria all'interno del cerchio si trasformano in teoremi della geometria di Lobačevskij. Viceversa, ogni teorema di quest'ultima trova un'interpretazione nella geometria ordinaria del cerchio.

Per esempio, l'assioma di Lobačevskij afferma che per un punto esterno a una retta passano almeno due rette che non la incontrano. Sostituendo alle rette le corde, otteniamo: per un punto interno al cerchio e non appartenente a una corda data passano almeno due corde che non incontrano la corda data. La verità di questa affermazione è evidente dalla figura 11. L'assioma di Lobačevskij è quindi soddisfatto.

Ricordiamo inoltre che fra le rette passanti per il punto dato e non incidenti alla retta data ve ne sono due estreme, dette parallele da Lobačevskij. Nel modello, fra le corde per A che non incontrano BC ve ne sono infatti due estreme: quelle dirette rispettivamente verso B e C. Esse non hanno punti comuni con BC, poiché i punti della circonferenza sono esclusi. Anche questo teorema è dunque soddisfatto.

Per tradurre altri teoremi occorre spiegare come misurare segmenti e angoli nel cerchio in modo corrispondente alla geometria di Lobačevskij. Le misure non possono essere quelle ordinarie: una corda ha lunghezza euclidea finita, mentre la retta che rappresenta è infinita. Potrebbe sembrare una contraddizione, ma vedremo che non lo è.

Ricordiamo come si misurano le lunghezze nella geometria ordinaria. Si sceglie un segmento *AB* come unità e si confronta con esso ogni altro segmento *XY*, riportando *AB* lungo *XY*. Se resta una parte più corta di *AB*, si divide l'unità, per esempio, in dieci parti uguali, cioè ottenibili l'una dall'altra mediante traslazione, e si riportano queste sul resto; se necessario, si passa ai centesimi e così via. La lunghezza viene così espressa da un numero decimale, eventualmente infinito. La misura si fonda dunque sul movimento dell'unità o di sue parti. Una volta definiti i movimenti, nel nostro caso come trasformazioni del cerchio che portano rette in rette, sappiamo quali segmenti considerare uguali e come misurare le lunghezze. La definizione di movimento contiene già implicitamente la regola di misura. Lo stesso vale per gli angoli, misurati riportando un angolo scelto come unità.

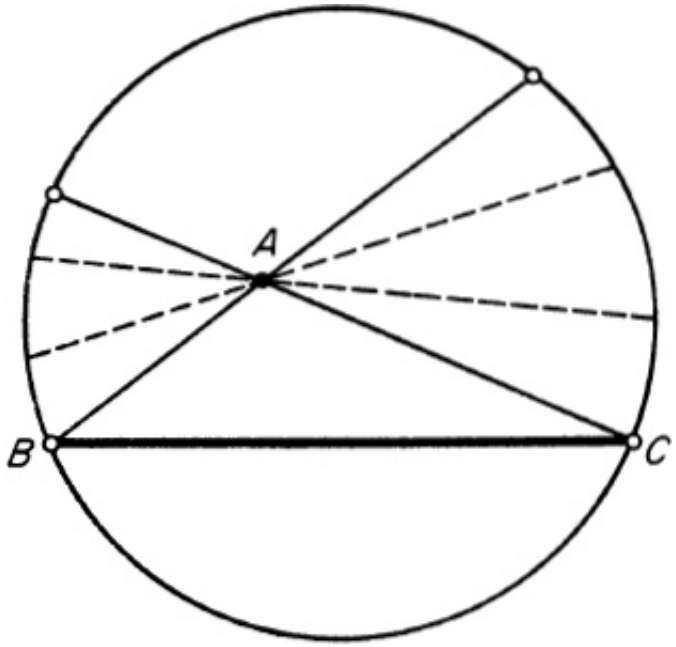


Figura 11.

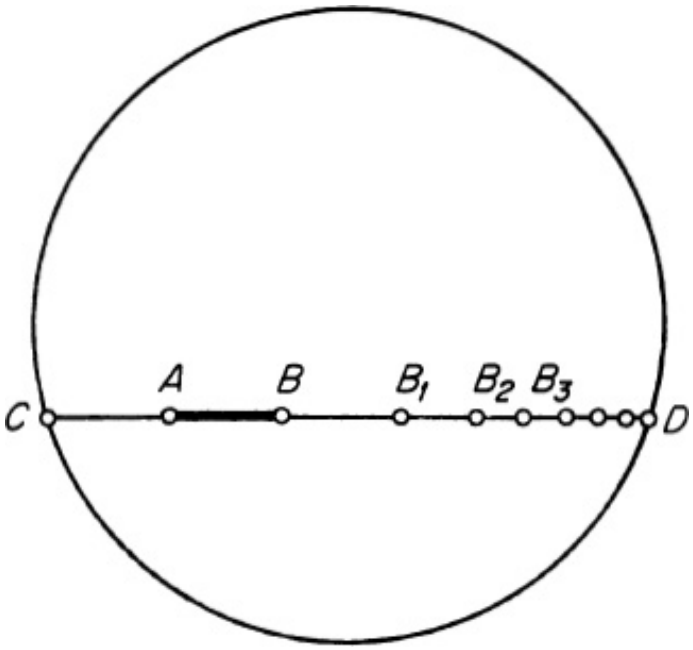


Figura 12.

Le regole corrispondenti nella geometria di Lobačevskij sono abbastanza semplici, benché sostanzialmente diverse da quelle ordinarie. Non le ricaveremo qui, poiché non sono essenziali per il nostro ragionamento.[9]

Secondo la nuova regola, una corda ha lunghezza infinita. Se un movimento del modello porta AB in BB_1, poi in B_1B_2 e così via, questi segmenti diventano sempre più corti nel senso euclideo, pur restando uguali nel modello (figura 12). I punti B_1, B_2, ..., B_k, ... si accumulano verso un estremo della corda. Ma quell'estremo è escluso dal nostro spazio: nel modello è «all'infinito». Nel senso della geometria di Lobačevskij, quei punti non si accumulano in alcun punto del piano, ma si allontanano indefinitamente. Riportando segmenti uguali mediante i movimenti definiti, non raggiungeremo mai dall'interno la circonferenza.

Per comprendere come si riportano i segmenti nel modello, consideriamo la trasformazione che svolge il ruolo di una traslazione lungo una retta.

Introduciamo coordinate cartesiane ortogonali con origine nel centro del cerchio. Supponiamo che il raggio sia uno: la circonferenza ha equazione $x^2 + y^2 = 1$ e i punti interni soddisfano $x^2 + y^2 < 1$.

Consideriamo la trasformazione definita dalle formule

$$x' = \frac{x + a}{1 + ax}, \quad y' = \frac{y\sqrt{1 - a^2}}{1 + ax}, \tag{3}$$

dove x', y' sono le coordinate dell'immagine del punto x, y e a è un numero assegnato arbitrariamente, con valore assoluto minore di 1.

Ricavando x e y dalle (3), si ottengono facilmente le formule inverse:

$$x = \frac{x' - a}{1 - ax'}, \quad y = \frac{y'\sqrt{1 - a^2}}{1 - ax'}. \tag{4}$$

La trasformazione (3) soddisfa le due condizioni richieste a un «movimento» del modello: porta il cerchio in sé e le rette in rette.

Per la prima proprietà dobbiamo verificare che la disuguaglianza o uguaglianza $x^2 + y^2 \leqslant 1$ comporti la relazione corrispondente $x'^2 + y'^2 \leqslant 1$, e viceversa. Mostriamo, per esempio, che da $x^2 + y^2 = 1$ segue $x'^2 + y'^2 = 1$: i punti della circonferenza rimangono dunque sulla circonferenza.

Calcoliamo $x'^2 + y'^2$ mediante le (3), tenendo conto che $y^2 = 1 - x^2$:

$$\begin{aligned} x'^2 + y'^2 &= \frac{(x + a)^2 + y^2(1 - a^2)}{(1 + ax)^2} = \frac{(x + a)^2 + (1 - x^2)(1 - a^2)}{(1 + ax)^2} \\ &= \frac{x^2 + 2ax + a^2 + 1 - x^2 - a^2 + a^2x^2}{(1 + ax)^2} = \frac{1 + 2ax + a^2x^2}{1 + 2ax + a^2x^2} \\ &= 1. \end{aligned}$$

Quindi, se $x^2 + y^2 = 1$, anche $x'^2 + y'^2 = 1$. Gli altri casi si verificano analogamente.

La seconda proprietà è altrettanto semplice. Ogni retta è rappresentata da un'equazione lineare e, viceversa, ogni equazione lineare non degenere rappresenta una retta. Sia data la retta

$$Ax + By + C = 0. \tag{5}$$

Sostituendo le espressioni (4), otteniamo

$$A\frac{x' - a}{1 - ax'} + B\frac{y'\sqrt{1 - a^2}}{1 - ax'} + C = 0,$$

oppure, riducendo al denominatore comune,

$$(A - aC)\,x' + B\sqrt{1 - a^2}\,y' + (C - aA) = 0.$$

L'equazione è lineare e rappresenta quindi una retta, precisamente l'immagine della (5). Osserviamo inoltre che la trasformazione (3) porta l'asse Ox in sé, spostandone soltanto i punti: su tale asse $y = 0$ e quindi anche $y' = 0$. La trasformazione sull'asse è descritta dalla sola formula

$$x' = \frac{x + a}{1 + ax} (|\,a\,| < 1). \tag{3'}$$

Il segmento di estremi x_1 e x_2 viene portato dalla (3′) nel segmento $x_1'x_2'$, che per convenzione consideriamo uguale al primo. Questo è il modo di «riportare un segmento» nel modello.

Per il centro O abbiamo $x = 0$, quindi $x' = a$: la (3′) porta il centro nel punto A di coordinata a.

Poiché a è arbitrario, purché $|a| < 1$, il centro può essere portato in qualunque punto del diametro sull'asse Ox.

La stessa trasformazione porta il punto inizialmente in A nel punto A_1 di coordinata

$$x_1 = \frac{a + a}{1 + a^2} = \frac{2a}{1 + a^2}.$$

Il segmento OA viene dunque trasformato in AA_1, cioè riportato sulla «retta» rappresentata dal diametro.

Ripetendo la trasformazione, possiamo riportare il segmento quante volte vogliamo. Il punto A_n, di coordinata x_n, viene portato in A_{n+1}, di coordinata

$$x_{n+1} = \frac{x_n + a}{1 + ax_n}.$$

Otteniamo così i punti $A, A_1, A_2, \ldots$, con coordinate

$$x_0 = a, \quad x_1 = \frac{2a}{1 + a^2}, \quad x_2 = \frac{x_1 + a}{1 + ax_1} = \frac{3a + a^3}{1 + 3a^2}, \cdots.$$

Tutti i segmenti A_nA_{n+1} si ottengono da OA mediante movimenti, quindi sono «uguali» nel senso della geometria rappresentata. È facile mostrare che, per a positivo, i punti A_n tendono all'estremo positivo del diametro; per a negativo tendono all'altro estremo. Nel modello si allontanano all'infinito.

Poiché possiamo orientare l'asse Ox in qualunque direzione, gli stessi spostamenti sono possibili lungo ogni diametro. Combinandoli con rotazioni attorno al centro e riflessioni rispetto a un diametro, otteniamo tutti i «movimenti» del modello. Nel prossimo paragrafo studieremo queste trasformazioni più precisamente e dimostreremo che il modello soddisfa davvero la geometria di Lobačevskij; in particolare, che i movimenti definiti soddisfano tutti gli assiomi geometrici del movimento.

Riassumiamo il modello di Klein. Lo spazio considerato è l'interno di un cerchio; i punti sono punti ordinari, le rette sono corde senza estremi, i movimenti sono trasformazioni biunivoche del cerchio aperto in sé che portano corde in corde. Le relazioni «un punto appartiene a una retta» e «un punto è fra altri due» mantengono il significato ordinario. Dalla definizione di movimento discendono le regole di misura di lunghezze, angoli e aree; sono definite anche l'uguaglianza dei segmenti, degli angoli e delle figure, e l'operazione di riportare un segmento su un altro.

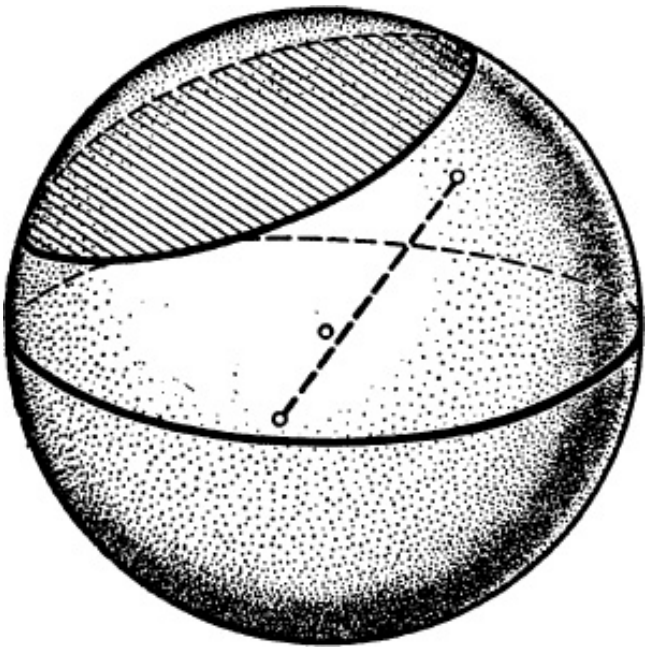


Figura 13.

Con queste convenzioni, a ogni teorema della geometria piana di Lobačevskij corrisponde un fatto della geometria euclidea nel cerchio; viceversa, ogni fatto espresso nei termini del modello può essere reinterpretato come teorema della geometria di Lobačevskij.

Analogamente si costruisce un modello della geometria spaziale. Come spazio si prende l'interno di una sfera (figura 13); le rette sono corde e i piani sono dischi ottenuti sezionando la sfera con piani euclidei. La superficie sferica è esclusa, e con essa gli estremi delle corde e le circonferenze dei dischi. I movimenti sono le trasformazioni biunivoche dell'interno della sfera in sé che portano corde in corde.

Il modello diede alla geometria di Lobačevskij un significato concreto semplice: essa può essere intesa come una particolare descrizione della geometria interna a un cerchio o a una sfera. Ne dimostrò contemporaneamente la coerenza relativa: i suoi risultati non possono produrre una contraddizione se è coerente la geometria euclidea, poiché ciascuno di essi si traduce nel linguaggio della geometria euclidea del modello.[10]

3. Altre interpretazioni.

Dopo Klein, il matematico francese Poincaré fornì un altro modello e lo impiegò per ottenere importanti risultati nella teoria delle funzioni di variabile complessa.[11] La geometria di Lobačevskij permise così di risolvere problemi difficili di un ramo completamente diverso della matematica. Essa ha trovato anche altre applicazioni matematiche e fisiche: per esempio, nel 1913 il fisico Varićak la applicò alla teoria della relatività.

La geometria di Lobačevskij continua a svilupparsi: vi si studiano le costruzioni geometriche, la teoria generale delle curve e delle superfici, i corpi convessi e molti altri argomenti.

§5. Gli assiomi della geometria e la loro verifica nel modello

1. Formulazione precisa degli assiomi.

Per dimostrare rigorosamente che il modello di Klein interpreta la geometria di Lobačevskij, dobbiamo precisare che cosa occorre dimostrare. Sarebbe assurdo verificarne i teoremi uno per uno: sono infiniti, perché se ne possono dedurre sempre di nuovi. Basta invece verificare le proposizioni fondamentali, dalle quali seguono tutte le altre. Ma queste devono essere formulate con precisione.

Il problema della coerenza della geometria di Lobačevskij conduce dunque a quello di enunciarne gli assiomi in modo preciso e completo. Poiché essi differiscono dagli assiomi euclidei soltanto per il postulato delle parallele, si tratta di formulare precisamente e completamente gli assiomi della geometria euclidea. In Euclide una tale formulazione non esiste ancora: mancano in particolare le proprietà del movimento o della sovrapposizione delle figure, benché siano usate nelle dimostrazioni. L'esigenza di completare e precisare gli assiomi emerse proprio con lo sviluppo della geometria non euclidea e con la tendenza generale, già ricordata, a rendere più rigorosi i fondamenti matematici alla fine dell'Ottocento.

Grazie alle ricerche di numerosi geometri, il problema della formulazione degli assiomi fu infine risolto.

Gli assiomi si possono scegliere in modi diversi, partendo da concetti differenti. Qui presenteremo quelli della geometria piana fondandoci sui concetti di punto, retta e movimento e sulle relazioni: il punto X appartiene alla retta a; B è fra A e C; un movimento porta X in Y. Gli altri concetti si definiscono mediante questi: per esempio, il segmento aperto fra due punti è l'insieme dei punti che si trovano fra essi; aggiungendo gli estremi si ottiene il segmento chiuso.

Gli assiomi si dividono in cinque gruppi.

I. Assiomi di incidenza.

1. Per due punti distinti passa una e una sola retta.

2. Ogni retta contiene almeno due punti distinti.

3. Esistono almeno tre punti non appartenenti a una stessa retta.

II. Assiomi di ordinamento.

1. Fra tre punti distinti di una retta, uno e uno soltanto si trova fra gli altri due.

2. Dati due punti distinti A e B di una retta, esiste sulla retta almeno un punto C tale che B si trovi fra A e C.

3. Una retta divide il piano in due semipiani: i punti esterni alla retta si dividono in due classi, tali che il segmento fra punti della stessa classe non incontri la retta, mentre quello fra punti di classi diverse la incontri.

III. Assiomi del movimento.

Un movimento va inteso come trasformazione dell'intero piano, non soltanto di una figura isolata.

1. Un movimento porta rette in rette.

2. Due movimenti eseguiti successivamente equivalgono a un unico movimento.

3. Siano A, A' due punti; a, a' due semirette uscenti da essi; α, α′ due semipiani delimitati dalle rette che contengono le rispettive semirette. Esiste un unico movimento che porta A in A', a in a' e α in α′. Intuitivamente, si porta prima A in A' con una traslazione, poi a in a' con una rotazione; infine, se i semipiani non coincidono, si esegue un ribaltamento attorno alla retta che contiene a'.

IV. Assioma di continuità.

1. Siano X_1, X_2, X_3, … punti di una retta, ciascuno a destra del precedente, ed esista un punto A a destra di tutti.[12] Allora esiste un punto B, anch'esso a destra di tutti gli X_n, al quale essi si avvicinano arbitrariamente: qualunque sia il punto C a sinistra di B, qualche X_n appartiene al segmento CB.

V. Assioma delle parallele di Euclide.

1. Per un punto esterno a una retta data passa una sola retta che non la incontra.

Questi assiomi bastano a costruire la geometria euclidea del piano. Da essi si possono dedurre tutte le proposizioni di un corso scolastico di geometria piana, anche se la deduzione completa è assai laboriosa.

Gli assiomi della geometria di Lobačevskij differiscono soltanto per quello delle parallele.

V′. Assioma delle parallele di Lobačevskij.

1. Per un punto esterno a una retta data passano almeno due rette che non la incontrano.

Può sembrare strano includere un assioma come «Ogni retta contiene almeno due punti». Nell'immagine abituale di una retta, i punti sono addirittura infiniti. Non sorprende che né Euclide né i matematici fino alla fine dell'Ottocento avessero sentito il bisogno di enunciarlo: veniva assunto tacitamente. Ora, però, la situazione è mutata. In una nuova interpretazione, «retta» può significare una geodetica su una superficie, una corda o qualcos'altro. Occorre quindi elencare precisamente ed esaustivamente tutto ciò che si richiede agli oggetti chiamati rette. Lo stesso vale per gli altri concetti e assiomi.

La comparsa di diverse interpretazioni geometriche fu dunque uno degli stimoli principali a precisarne i fondamenti. Anche storicamente, la formulazione rigorosa degli assiomi seguì i modelli di Beltrami, Klein e Poincaré.

2. Verifica degli assiomi nel modello di Klein.

Mostriamo ora che nel modello sono soddisfatti tutti gli assiomi elencati, salvo quello euclideo delle parallele. Come abbiamo osservato nel paragrafo precedente (figura 11), al suo posto vale l'assioma di Lobačevskij. Restano da verificare i gruppi I-IV.

Il piano è l'interno di un cerchio di raggio uno. I punti sono quelli ordinari e le rette sono corde; appartenenza e interposizione conservano il significato usuale. È quindi evidente che valgono gli assiomi di incidenza, ordinamento e continuità. Il terzo assioma di ordinamento, per esempio, afferma semplicemente che una corda divide il cerchio in due parti.

Restano gli assiomi del movimento. Per definizione, un movimento porta il cerchio in sé e le rette in rette. I primi due assiomi sono immediati: il primo perché le rette sono corde e queste vengono conservate; il secondo perché la composizione di due trasformazioni che portano il cerchio in sé e le corde in corde possiede le stesse proprietà ed è quindi ancora un movimento.

Rimane soltanto il terzo assioma: la sua verifica è l'unica parte difficile.

Osserviamo anzitutto che esso contiene due affermazioni.

Siano A, A' due punti, a, a' due semirette uscenti da essi e α, α' due semipiani delimitati dalle rette che le contengono.

La prima affermazione è che esiste un movimento che porta A in A', a in a' e α in α'.

La seconda è che tale movimento è unico.

Potremmo richiamare la dimostrazione già data nel capitolo III, §14, ma preferiamo ripeterla qui senza collegarla, come allora, a problemi più generali.

Verifichiamo la prima affermazione nel modello, interpretando opportunamente «semiretta», «semipiano» e «movimento».

Supponiamo dapprima che A' sia il centro del cerchio. Scegliamo assi con origine nel centro e asse Ox passante per A (figura 14). Nel paragrafo precedente abbiamo studiato la trasformazione

$$x' = \frac{x + a}{1 + ax}, \quad y' = \frac{y\sqrt{1 - a^2}}{1 + ax}. \tag{6}$$

e abbiamo dimostrato che è un «movimento»: porta il cerchio in sé e le rette in rette.

Sia x_0 l'ascissa di A, la cui ordinata è $y_0 = 0$. Scegliendo $a = -x_0$, la (6) porta A nel punto (0, 0), cioè in A'.

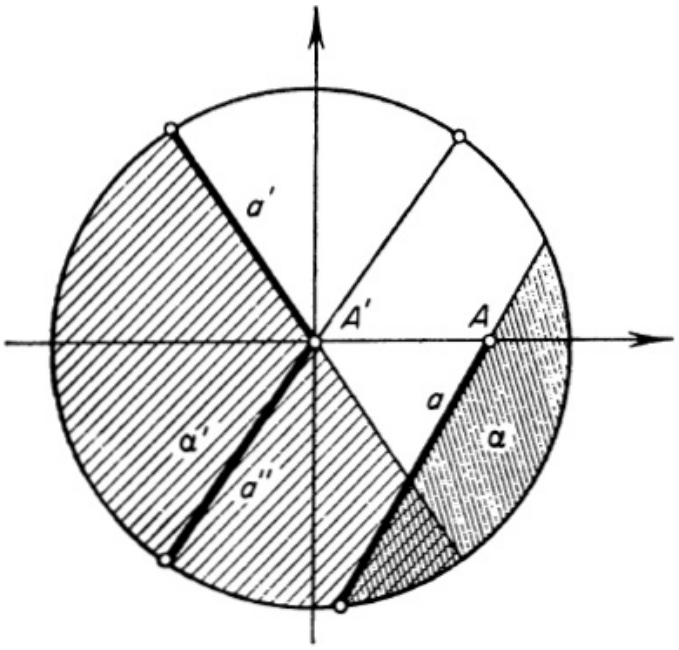


Figura 14.

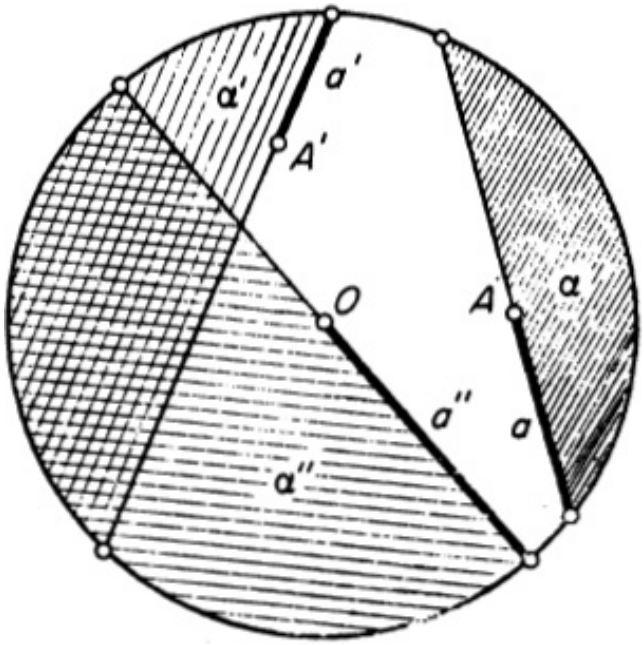


Figura 15.

Poiché le rette vengono trasformate in rette, la «semiretta» a, ossia una parte di corda, assume una posizione a'' (figura 14). Una rotazione attorno al centro porta poi a'' in a'. Il «semipiano» α è una delle due parti del cerchio delimitate dalla corda che contiene a. Se la sua immagine coincide già con α′, abbiamo finito; altrimenti, una riflessione rispetto al diametro che contiene a' la porta nel semicerchio α′.

Combinando lo spostamento (6), una rotazione e, se necessario, una riflessione, abbiamo dunque portato A, a, α in A', a', α′. La composizione di questi movimenti è ancora un movimento: l'esistenza richiesta è dimostrata.

Finora abbiamo supposto A' nel centro. Se si trova altrove, quanto appena dimostrato permette di portarlo nel centro mediante un movimento D_1. La semiretta a' diventa allora una semiretta a'' uscente dal centro e il semipiano α' diventa un semicerchio α'' (figura 15). Un altro movimento D_2 porta A nel centro, a in a'' e α in α''. Applicando infine l'inverso di D_1, il centro viene riportato in A', mentre a'' e α'' tornano rispettivamente in a' e α'.[13]

La composizione di D_2 con l'inverso di D_1 porta dunque A, a, α in A', a', α'. Essendo ancora un movimento, essa dimostra l'esistenza richiesta per qualunque posizione dei due punti nel cerchio. La prima parte del terzo assioma è verificata in tutta la sua generalità.

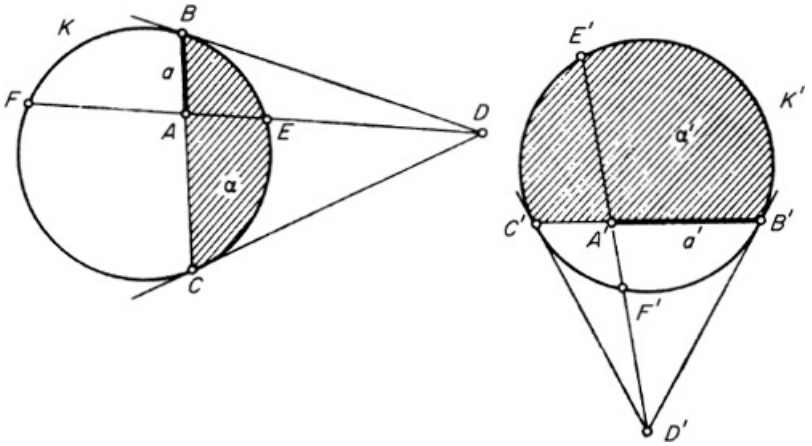


Figura 16.

Dimostriamo ora la seconda parte, cioè l'unicità.

Siano A, A' punti interni al cerchio; a, a' parti di corde uscenti da essi; α, α' le parti del cerchio delimitate dalle rette corrispondenti. Per chiarezza li rappresentiamo in due disegni distinti (figura 16), benché appartengano allo stesso cerchio. Affermiamo che il movimento che porta rispettivamente A, a, α in A', a', α' è unico, cioè completamente determinato da questi dati.

Per dimostrarlo consideriamo l'estensione della trasformazione dal cerchio all'intero piano proiettivo.[14] Un movimento, per definizione, porta rette in rette; la corrispondente trasformazione è proiettiva. I nostri movimenti sono quindi trasformazioni proiettive che conservano il cerchio. Nella figura 16 rappresentiamo il cerchio K e la sua immagine K' separatamente: basta immaginare di sovrapporli mediante una traslazione.

Nel capitolo III, §13, abbiamo studiato le trasformazioni proiettive e dimostrato il seguente teorema, che useremo qui: una trasformazione proiettiva è completamente determinata dalle immagini di quattro punti, nessuna terna dei quali sia allineata.

Il movimento considerato porta la parte di corda a in a', quindi l'estremo B in B'. Poiché conserva le corde, porta anche l'altro estremo C in C'.

Inoltre, poiché porta rette in rette e il cerchio K nel cerchio K', le tangenti in B e C vengono portate nelle tangenti in B' e C'. Il loro punto d'incontro D va quindi nel punto d'incontro D' delle immagini (figura 16).[15]

Poiché A va in A', la retta AD va in $A'D'$. La prima incontra la circonferenza in E, F, la seconda in E', F'. La trasformazione porta dunque la prima coppia nella seconda. Siano E sull'arco che delimita α ed E' sull'arco che delimita α′. Poiché α va in α′, necessariamente E va proprio in E' e F in F'.

Abbiamo così determinato le immagini dei quattro punti B, C, E, F: sono rispettivamente B', C', E', F'. In ciascuna quaterna nessuna terna è allineata, poiché i punti appartengono a una circonferenza. Il teorema citato garantisce allora l'unicità della trasformazione proiettiva e quindi del movimento richiesto.

Nel modello sono dunque soddisfatti tutti gli assiomi euclidei tranne quello delle parallele, sostituito dall'assioma di Lobačevskij. Il modello realizza quindi la geometria non euclidea, riconducendola alla geometria euclidea interna al cerchio con il particolare significato assegnato a «retta» e «movimento». Possiamo anche sviluppare direttamente la geometria di Lobačevskij mediante questo modello concreto, un metodo spesso più comodo per risolvere i problemi.

Dal punto di vista dell'analisi logica dei fondamenti, la dimostrazione stabilisce due fatti: la geometria di Lobačevskij è coerente se lo è quella euclidea; il postulato euclideo delle parallele non può essere dedotto dagli altri assiomi elencati.

§6. Teorie geometriche autonome nate dalla geometria euclidea

1. Geometria proiettiva.

Accanto alla creazione della geometria di Lobačevskij, un altro sviluppo fondamentale trasformò la disciplina. Nella grande varietà delle proprietà geometriche dello spazio si isolarono gruppi di proprietà caratterizzati da legami particolari e da una propria stabilità, studiandoli autonomamente. Con metodi specifici, queste ricerche diedero origine a nuovi rami della scienza delle forme spaziali, così come anatomia e fisiologia sono rami distinti della scienza dell'organismo umano.

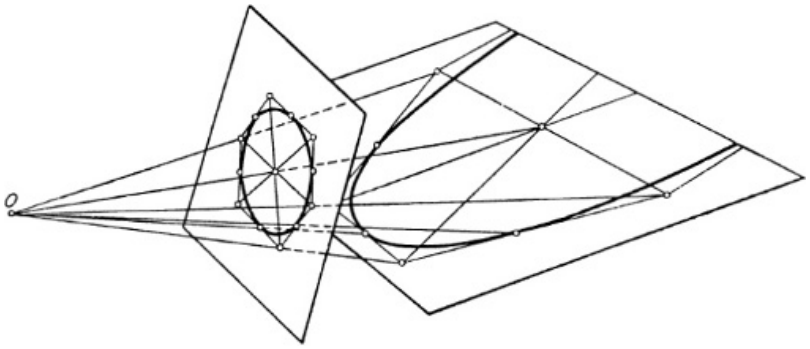


Figura 17.

In origine la geometria non era suddivisa. Studiava soprattutto le proprietà metriche, legate alla misura delle dimensioni delle figure. Le proprietà qualitative della loro disposizione, indipendenti dalla misura, venivano considerate solo incidentalmente. Eppure si era osservato da tempo che alcune di esse restano invariate anche quando forma e posizione subiscono alterazioni considerevoli.

Consideriamo, per esempio, la proiezione di una figura da un piano a un altro (figura 17). Le lunghezze e gli angoli cambiano e i contorni risultano deformati; restano però conservate proprietà quali l'allineamento di più punti o la tangenza di una retta a una curva.

Nel capitolo III abbiamo già trattato proiezioni e trasformazioni proiettive, collegandole alla prospettiva, cioè alla rappresentazione sul piano delle figure spaziali. Lo studio della prospettiva risale agli antichi architetti e a Euclide. Se ne occuparono artisti come Dürer e Leonardo da Vinci e, nel Seicento, l'ingegnere e matematico Desargues. All'inizio dell'Ottocento Poncelet isolò e studiò sistematicamente le proprietà conservate da tutte le trasformazioni proiettive del piano o dello spazio, creando così una disciplina autonoma: la geometria proiettiva.[16]

Potrebbe sembrare che le proprietà conservate da tutte le trasformazioni proiettive siano poche ed elementari, ma non è così. Per esempio, il teorema secondo cui i punti d'intersezione dei lati opposti, prolungati, di un esagono inscritto in una circonferenza sono allineati vale anche per ellissi, parabole e iperboli. Non è immediato accorgersene: il teorema riguarda

proprietà proiettive e queste curve si ottengono dalla circonferenza per proiezione. Ancora meno evidente è il legame con il teorema secondo cui le tre diagonali principali di un esagono circoscritto si incontrano in un punto: soltanto la geometria proiettiva ne rivela la profonda analogia. Né è ovvio che, nonostante la deformazione delle distanze, per quattro punti allineati A, B, C, D resti invariato il birapporto $AC/CB : AD/DB$ (figura 18):

$$\frac{AC}{CB} : \frac{AD}{DB} = \frac{A'C'}{C'B'} : \frac{A'D'}{D'B'} .$$

La prospettiva conserva quindi molte relazioni. Per esempio, conoscendo la distanza reciproca dei pali telegrafici A, B, C, questa proprietà permette di determinarne la distanza dal punto D su una fotografia della strada che si allontana (figura 18).

Abbiamo già discusso nel capitolo III la geometria proiettiva e le sue applicazioni alla fotografia aerea. Le sue leggi trovano naturalmente impiego anche in architettura, nella costruzione di panorami, nella scenografia e nella decorazione.

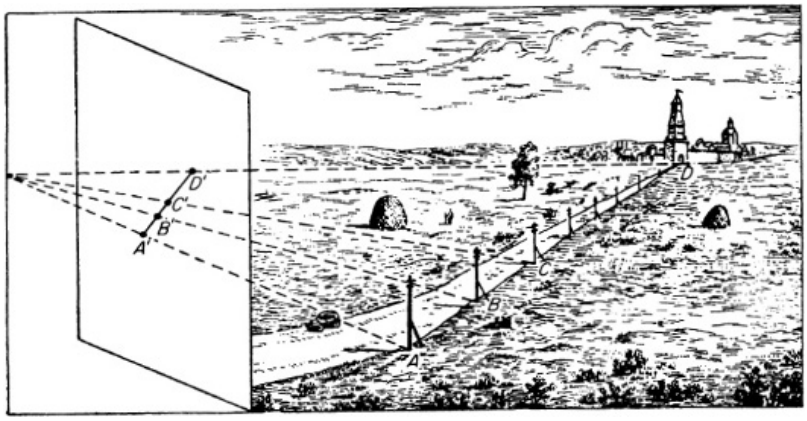


Figura 18.

La costituzione della geometria proiettiva come disciplina autonoma ebbe un ruolo importante nello sviluppo della geometria stessa.

2. Geometria affine.

Un altro esempio è la geometria affine. Essa studia le proprietà che non cambiano sotto trasformazioni nelle quali le coordinate cartesiane iniziali (x, y, z) e finali (x', y', z') di ogni punto sono legate da equazioni lineari:

$$\begin{aligned} x' &= a_1x + b_1y + c_1z + d_1 , \\ y' &= a_2x + b_2y + c_2z + d_2 , \\ z' &= a_3x + b_3y + c_3z + d_3 \end{aligned}$$

Si suppone che il determinante

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

sia diverso da zero.

Ogni trasformazione affine si può ricondurre a un movimento, eventualmente comprendente una riflessione rispetto a un piano, seguito da dilatazioni o contrazioni lungo tre direzioni fra loro perpendicolari.

Queste trasformazioni conservano numerose proprietà. Le rette restano rette, e in effetti sono conservate tutte le proprietà proiettive; inoltre le parallele restano parallele. Restano invariati i rapporti fra volumi, fra aree di figure poste nello stesso piano o in piani paralleli e fra lunghezze di segmenti sulla stessa retta o su rette parallele. Molti teoremi noti appartengono essenzialmente alla geometria affine: le mediane di un triangolo sono concorrenti; le diagonali di un parallelogramma si dimezzano reciprocamente; i punti medi di corde parallele di un'ellisse sono allineati.

Tutta la teoria delle curve e delle superfici di secondo ordine è strettamente legata alla geometria affine. La distinzione stessa fra ellissi, parabole e iperboli si fonda su proprietà affini: una trasformazione affine porta un'ellisse in un'ellisse, mai in una parabola o in un'iperbole. Analogamente una parabola può essere trasformata in qualunque altra parabola, ma non in un'ellisse, e così via.

L'importanza dello studio autonomo delle proprietà affini è rafforzata dal fatto che trasformazioni molto più complesse, se differenziabili con differenziale non singolare, sono approssimate localmente da trasformazioni affini. Proprio lo studio di regioni infinitesime è alla base delle applicazioni del calcolo differenziale.

3. Il programma di Erlangen di Klein.

Nel 1872, in una relazione all'Università di Erlangen oggi nota come *Programma di Erlangen*, Klein riassunse lo sviluppo delle geometrie proiettiva, affine e di altri tipi, formulandone chiaramente il principio generale: si può considerare un gruppo arbitrario di trasformazioni biunivoche dello spazio e studiare le proprietà delle figure conservate da tutte le trasformazioni del gruppo.[17]

Da questo punto di vista le proprietà spaziali si dispongono, per così dire, in strati secondo la loro profondità e stabilità. La geometria euclidea nasce astraendo da tutte le proprietà dei corpi reali tranne quelle geometriche. Nei suoi rami particolari compiamo un'ulteriore astrazione, trascurando tutte le proprietà geometriche salvo quelle che interessano la teoria considerata.

Il principio di Klein permette di costruire molte geometrie. Possiamo, per esempio, considerare le trasformazioni che conservano gli angoli fra curve, dette conformi, e studiare le

proprietà da esse conservate: otteniamo una geometria conforme. Non è necessario trasformare l'intero spazio. Considerando soltanto punti e corde di un cerchio e tutte le trasformazioni del cerchio in sé che conservano le corde, otteniamo la geometria che nei §§4 e 5 abbiamo riconosciuto come geometria di Lobačevskij.

4. Spazi astratti.

Lo sviluppo delle teorie così isolate non si arrestò a questi risultati, neppure dal punto di vista teorico, oltre che per la ricchezza del loro contenuto.

Se ci interessano soltanto le proprietà affini, possiamo astrarre da tutte le altre e immaginare uno spazio e figure che possiedano soltanto quelle proprietà. In questo «spazio» non sono assegnate proprietà ulteriori. È naturale cercare una descrizione assiomatica della sua geometria: considerare oggetti astratti chiamati «punti», «rette» e «piani», con proprietà espresse da assiomi, meno numerose di quelle euclidee, in modo che le conseguenze corrispondano alle proprietà affini delle figure ordinarie.

Questo è effettivamente possibile. Una tale collezione di «punti», «rette» e «piani», con il sistema delle loro relazioni, si chiama spazio affine.

Allo stesso modo possiamo immaginare un sistema astratto di oggetti che possiedano soltanto le proprietà proiettive delle figure euclidee. In questo caso la differenza fra il sistema di assiomi e quello della geometria ordinaria è ancora più profonda.

5. Spazio topologico astratto.

Penetrando più a fondo nella natura delle forme geometriche, si incontrano proprietà ancora più generali di quelle proiettive. Esse sono tanto strettamente legate alla figura da conservarsi sotto deformazioni arbitrarie, purché non si producano strappi o incollamenti fra sue parti. Per precisare questa descrizione intuitiva ricorriamo alla continuità, nota dall'analisi: consideriamo una trasformazione biunivoca dei punti della figura in una nuova posizione, tale che le nuove coordinate siano funzioni continue delle vecchie e, viceversa, le vecchie coordinate siano funzioni continue delle nuove.

Le proprietà conservate da tutte queste trasformazioni si chiamano topologiche; la disciplina che le studia è la topologia (si veda il capitolo XVIII).

Nei casi più semplici, le proprietà topologiche hanno un carattere particolarmente intuitivo. Sembra quasi evidente, per esempio, che una curva piana ottenuta deformando continuamente una circonferenza senza incollarne punti distinti divida il piano in una regione interna e una esterna, per quanto tortuoso sia il suo contorno. La proprietà di separare il piano è dunque topologica. È anche intuitivamente chiaro che la superficie di un toro (figura 19) non può trasformarsi in una sfera mediante una trasformazione continua biunivoca con inversa continua. La possibilità di trasformare una superficie in un toro è quindi una proprietà topologica che la distingue da molte altre superfici.

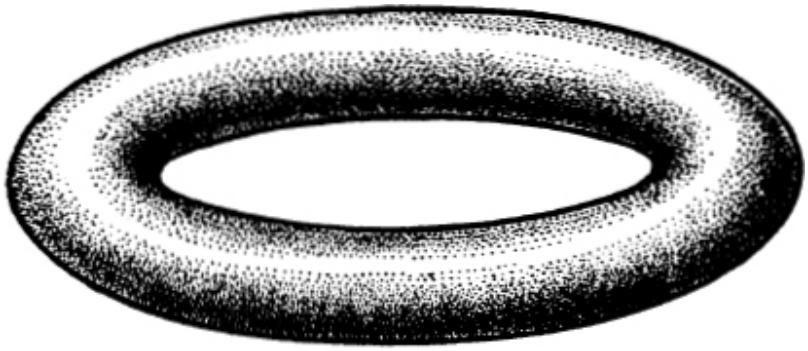


Figura 19.

I ragionamenti fondati sulla continuità sono intuitivi e spesso chiariscono così bene il problema da suggerire di trasformarli in dimostrazioni rigorose e di estenderli a questioni molto più complesse.

Consideriamo, per esempio, il ragionamento che conduce al teorema fondamentale dell'algebra: ogni equazione polinomiale non costante

$$z^n + a_1 z^{n-1} + a_2 z^{n-2} + \cdots + a_{n-1} z + a_n = 0 \qquad (7)$$

possiede almeno una radice reale o complessa.

Sia z un punto di un piano complesso e $w = f(z)$ il punto corrispondente di un secondo piano, dove f è il primo membro della (7). Per $|z|$ molto grande, $f(z)$ differisce relativamente poco da z^n. Quest'ultima funzione è semplice: quando z percorre una circonferenza centrata nell'origine, $w_1 = z^n$ percorre esattamente n volte una circonferenza di raggio $|z|^n$ nel piano w.[18] Perciò $w = f(z)$ descrive un contorno Γ vicino a quella circonferenza e avvolto n volte attorno all'origine (figura 20).

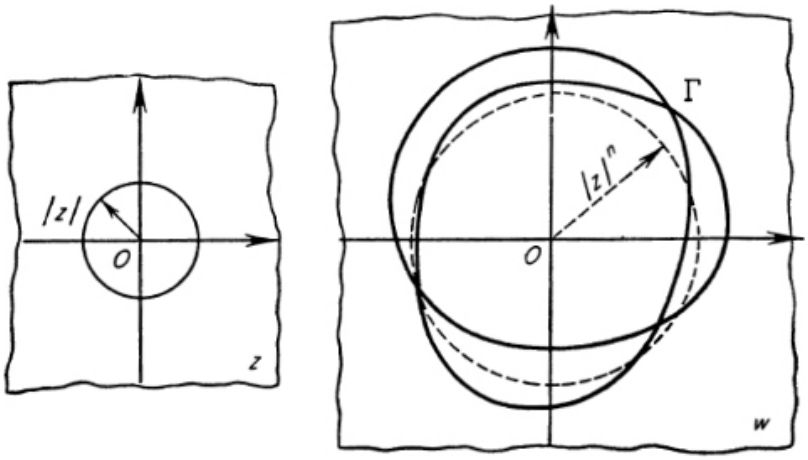


Figura 20.

Contraendo continuamente a un punto la circonferenza descritta da z, anche il contorno Γ descritto da $f(z)$ si deforma continuamente e si contrae a un punto. Ma un contorno inizialmente avvolto attorno all'origine non può contrarsi a un punto evitando sempre l'origine. Deve quindi passarvi almeno una volta: per qualche z si ha $w = f(z) = 0$, e questo z è

una radice della (7). Il numero iniziale di avvolgimenti suggerisce inoltre che le radici siano esattamente n, contate con la loro molteplicità.

Naturalmente il ragionamento richiede di stabilire rigorosamente le proprietà topologiche del contorno e delle sue deformazioni che abbiamo utilizzato.

Potremmo citare molti altri esempi dell'uso di proprietà topologiche in rami della matematica spesso assai lontani dalla geometria.

Anche qui possiamo immaginare un insieme astratto di oggetti dotato soltanto di proprietà di questo genere (si veda il capitolo XX, §7). Un tale insieme si chiama spazio topologico astratto.

Questa prospettiva è molto più ampia dello studio delle sole figure geometriche. Gli spazi topologici sono estremamente vari: i punti della superficie di un toro, con le loro relazioni di vicinanza, formano uno spazio; i punti del piano ne formano un altro; lo spazio euclideo intero un terzo. Anche le superfici di Riemann a più fogli, incontrate nel capitolo IX, §5, sono spazi topologici. Ma il fatto più notevole è che si possono definire chiaramente vicinanza e prossimità anche fra oggetti che non considereremmo punti geometrici. Per tutte le posizioni possibili di un meccanismo articolato, per esempio, sappiamo indicare che cosa significhi una posizione «vicina» a un'altra, oppure una famiglia di posizioni contenente elementi arbitrariamente vicini a una posizione data.

Il concetto di spazio topologico è dunque estremamente generale. Vi torneremo nel §8.

Questo paragrafo non voleva soltanto presentare diverse geometrie, ma mostrare come problemi concreti conducano a isolare e studiare gruppi particolari di proprietà geometriche. Lo studio di queste proprietà nella loro forma pura suggerisce poi oggetti astratti che possiedono soltanto quelle proprietà, cioè il corrispondente spazio astratto.

Nel prossimo paragrafo esamineremo altri sviluppi, che conducono a spazi astratti di natura diversa.

§7. Spazi multidimensionali

1. Geometria dello spazio a *n* dimensioni.

Un passo importante nello sviluppo delle nuove idee geometriche fu la creazione della geometria multidimensionale, già ricordata nel capitolo precedente. Fra le sue motivazioni vi era il desiderio di usare ragionamenti geometrici per risolvere problemi di algebra e analisi. Questo approccio si fonda sul metodo delle coordinate. Consideriamo un esempio semplice.

Vogliamo sapere quante soluzioni intere abbia la disuguaglianza $x^2 + y^2 < N$. Interpretando x, y come coordinate cartesiane, la domanda diventa: quanti punti a coordinate intere si trovano all'interno di un cerchio di raggio $\sqrt{N}$? Tali punti sono i vertici dei quadrati unitari che ricoprono il piano (figura 21). Il numero di punti interni al cerchio è approssimativamente uguale al numero di quadrati che esso contiene, cioè all'area del cerchio di raggio $\sqrt{N}$. Il numero cercato è quindi circa πN. Non è difficile dimostrare che l'errore relativo tende a zero per $N \to \infty$. Determinare più precisamente questo errore è invece un problema difficile della teoria dei numeri, oggetto di ricerche profonde in epoca relativamente recente rispetto alla stesura del libro.

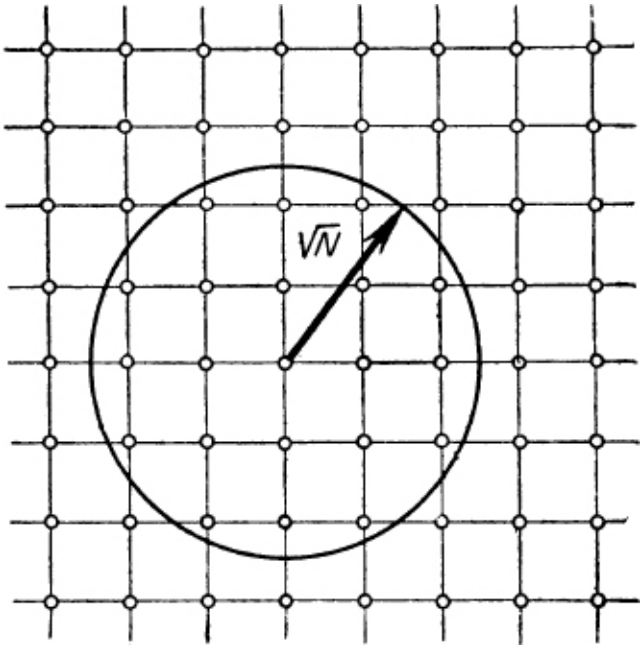


Figura 21.

È bastato tradurre il problema in linguaggio geometrico per ottenere un risultato tutt'altro che evidente dal punto di vista puramente algebrico. Lo stesso metodo vale per una disuguaglianza in tre incognite. Oltre tre incognite, però, lo spazio ordinario non basta, perché un suo punto è individuato da tre coordinate. Per conservare l'utile analogia introduciamo uno spazio astratto a n dimensioni, i cui punti siano determinati da n coordinate $x_1, \ldots, x_n$. Generalizziamo poi i concetti geometrici in modo da applicare ragionamenti analoghi ai problemi con n variabili, facilitandone la soluzione. La generalizzazione è possibile grazie

all'uniformità delle leggi algebriche, che permettono di risolvere molti problemi per un numero arbitrario di variabili. I ragionamenti geometrici validi in tre dimensioni possono così estendersi a un numero qualsiasi di dimensioni.

2. Coordinate nella geometria a *n* dimensioni.

Tracce dell'idea di uno spazio quadridimensionale si trovano già in Lagrange, che nei lavori di meccanica considerava formalmente il tempo come «quarta coordinata» accanto alle tre spaziali. La prima trattazione sistematica degli elementi della geometria multidimensionale fu però data nel 1844 da Grassmann e, indipendentemente, da Cayley. Essi partirono dall'analogia formale con la geometria analitica ordinaria. Eccone lo schema, in una formulazione moderna.

Un punto dello spazio a n dimensioni è determinato da n coordinate $x_1, \ldots, x_n$. Una figura è un luogo geometrico, ossia un insieme di punti che soddisfano certe condizioni. Un «cubo» a n dimensioni, per esempio, è l'insieme dei punti le cui coordinate verificano $a \leqslant x_i \leqslant b$ per $i = 1, 2, \ldots, n$. L'analogia è evidente: per $n = 3$, queste disuguaglianze definiscono un cubo di spigolo $b - a$, con spigoli paralleli agli assi. La figura 22 rappresenta il caso $a = 0$, $b = 1$.

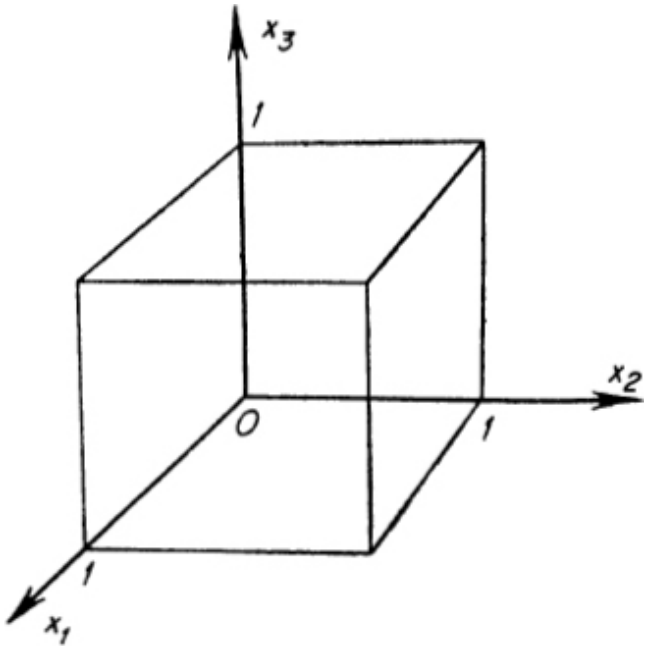


Figura 22.

La distanza fra due punti si definisce come la radice quadrata della somma dei quadrati delle differenze fra le coordinate:

$$d = \sqrt{(x_1' - x_1'')^2 + (x_2' - x_2'')^2 + \cdots + (x_n' - x_n'')^2}.$$

È la generalizzazione diretta delle formule note per la distanza nel piano o nello spazio tridimensionale, cioè per $n = 2$ o 3.

Possiamo ora definire la congruenza delle figure: due figure sono congruenti se esiste una corrispondenza biunivoca fra i loro punti che conservi le distanze fra ogni coppia. Una

trasformazione che conserva le distanze si chiama movimento generalizzato.[19] Per analogia con il caso ordinario, la geometria a *n* dimensioni studia le proprietà delle figure conservate da questi movimenti. Questa definizione, formulata negli anni Settanta dell'Ottocento, diede un fondamento preciso alla teoria. Da allora la geometria multidimensionale è stata studiata in direzioni analoghe a quelle della geometria euclidea: geometria elementare, teoria generale delle curve e così via.

La distanza permette di estendere altri concetti: segmento, sfera, lunghezza, angolo, volume. Per esempio, una sfera piena, o palla, a *n* dimensioni è l'insieme dei punti la cui distanza da un centro dato non supera un numero *R* assegnato. È dunque descritta dalla disuguaglianza

$$(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2 \leqslant R^2,$$

dove a_1, ..., a_n sono le coordinate del centro. La superficie sferica è definita dall'equazione

$$(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2 = R^2.$$

Il segmento *AB* si può definire come l'insieme dei punti *X* per i quali la somma delle distanze da *A* e da *B* è uguale alla distanza fra *A* e *B*. La lunghezza del segmento è la distanza fra i suoi estremi.

3. Piani e iperpiani.

Soffermiamoci sui piani di varie dimensioni.

Nello spazio tridimensionale troviamo i «piani» unidimensionali, cioè le rette, e i piani ordinari, bidimensionali. In uno spazio a *n* dimensioni, con $n > 3$, dobbiamo considerare anche piani di dimensione da 3 a $n - 1$.

Nello spazio tridimensionale un piano è rappresentato da un'equazione lineare, una retta da due equazioni lineari indipendenti.

Generalizzando direttamente, definiamo un piano a *k* dimensioni nello spazio a *n* dimensioni come il luogo dei punti le cui coordinate soddisfano un sistema di $n - k$ equazioni lineari:

$$\begin{gathered} a_{11}x_1 + \ldots + a_{1n}x_n + b_1 = 0, \\ a_{21}x_1 + \ldots + a_{2n}x_n + b_2 = 0, \\ \ldots \\ a_{n-k,1}x_1 + \ldots + a_{n-k,n}x_n + b_{n-k} = 0. \quad (8) \end{gathered}$$

Le equazioni devono essere compatibili e indipendenti, cioè nessuna deve essere conseguenza delle altre. Ciascuna rappresenta un iperpiano di dimensione $n - 1$; insieme ne determinano i punti comuni. La compatibilità significa che esistono punti comuni agli $n - k$ iperpiani. L'indipendenza significa che nessuna equazione si può omettere: altrimenti il sistema si ridurrebbe a meno equazioni e definirebbe un piano di dimensione maggiore. Geometricamente, un piano di dimensione *k* è dunque l'intersezione di $n - k$ iperpiani definiti da equazioni indipendenti. Per $k = 1$ abbiamo $n - 1$ equazioni e otteniamo una retta. È una generalizzazione naturale dei risultati della geometria analitica. Il suo vantaggio è rendere più

chiare le proprietà dei sistemi lineari interpretandole geometricamente, come abbiamo già fatto nel capitolo XVI. L'originale usa «iperpiano» anche per piani di dimensione inferiore a $n - 1$; qui riserviamo il termine a quelli di dimensione $n - 1$.

Una proprietà importante di un piano di dimensione k è che esso può essere considerato a sua volta uno spazio a k dimensioni. Un piano tridimensionale, per esempio, è uno spazio tridimensionale ordinario. Ciò permette di trasferire molti risultati da dimensioni inferiori a dimensioni superiori, mediante il consueto passaggio induttivo da n a $n + 1$.

L'algebra dimostra che, se le (8) sono compatibili e indipendenti, si possono scegliere opportunamente k variabili fra le n, assegnare loro valori arbitrari ed esprimere mediante esse le altre $n - k$ variabili.[20] Per esempio:

$$x_{k+1} = c_{11}x_1 + \ldots + c_{1k}x_k + d_1,$$
$$\ldots$$
$$x_n = c_{n-k,1}x_1 + \ldots + c_{n-k,k}x_k + d_{n-k}.$$

Qui x_1, x_2, ..., x_k possono assumere valori arbitrari, mentre le altre coordinate sono determinate da esse. La posizione di un punto del piano è quindi individuata da k coordinate indipendenti: è in questo senso che il piano ha k dimensioni.

Dalla definizione dei piani si ricavano per via puramente algebrica i seguenti teoremi fondamentali.

1. Per $k + 1$ punti che non appartengono a uno stesso piano di dimensione $k - 1$ passa uno e un solo piano di dimensione k.

L'analogia con i fatti elementari è evidente. La dimostrazione usa la teoria dei sistemi lineari ed è piuttosto laboriosa, perciò non la riportiamo.

2. Se un piano di dimensione l e uno di dimensione k, in uno spazio a n dimensioni, hanno almeno un punto comune e $l + k \geqslant n$, la loro intersezione è un piano di dimensione non inferiore a $l + k - n$.

Ne segue, come caso particolare, che due piani bidimensionali nello spazio tridimensionale, distinti e non paralleli, si incontrano lungo una retta: $n = 3$, $l = k = 2$, quindi $l + k - n = 1$. In quattro dimensioni, invece, due piani bidimensionali possono avere un solo punto comune. Per esempio, i piani definiti dai sistemi

$$\left.\begin{aligned} x_1 &= 0 \\ x_2 &= 0 \end{aligned}\right\}, \qquad \left.\begin{aligned} x_3 &= 0 \\ x_4 &= 0 \end{aligned}\right\},$$

si incontrano soltanto nel punto di coordinate $x_1 = x_2 = x_3 = x_4 = 0$.

La dimostrazione è semplicissima. Il piano di dimensione l è definito da $n - l$ equazioni, quello di dimensione k da $n - k$. I punti comuni soddisfano tutte le $(n - l) + (n - k) = n - (l + k - n)$ equazioni. Se esse sono indipendenti, l'intersezione ha per definizione dimensione $l + k - n$; altrimenti ha dimensione maggiore.

A questi due teoremi ne aggiungiamo altri due.

3. In ogni piano di dimensione k esistono almeno $k + 1$ punti che non appartengono a un piano di dimensione inferiore. In uno spazio a n dimensioni esistono almeno $n + 1$ punti non contenuti in alcun iperpiano.

4. Se una retta ha due punti distinti in comune con un piano di dimensione qualsiasi, è interamente contenuta in esso. Più generalmente, se un piano di dimensione l ha in comune con un piano di dimensione k almeno $l + 1$ punti non appartenenti a un piano di dimensione $l - 1$, allora è contenuto nel secondo.

La geometria a n dimensioni può essere costruita anche partendo da assiomi che generalizzano quelli del §5. In tal caso i quattro teoremi appena enunciati svolgono il ruolo di assiomi di incidenza. Il concetto di assioma è dunque relativo alla costruzione scelta: una stessa proposizione può essere un teorema in una presentazione e un assioma in un'altra.

4. Esempi di spazi a *n* dimensioni.

Abbiamo ora un'idea generale del concetto matematico di spazio multidimensionale. Per chiarirne il significato fisico, torniamo alla rappresentazione grafica. Se vogliamo rappresentare la dipendenza della pressione di un gas dal volume, riportiamo su due assi il volume v e la pressione p. Nelle condizioni assegnate, la dipendenza è rappresentata da una curva: per un gas ideale a temperatura costante, la legge di Boyle-Mariotte dà un'iperbole. Ma se lo stato di un sistema più complesso è determinato da cinque dati anziché da due, la sua rappresentazione grafica conduce a uno spazio a cinque dimensioni.

Consideriamo una lega di tre metalli o una miscela di tre gas. Nel modello considerato, lo stato è individuato da quattro dati: temperatura T, pressione p e percentuali c_1, c_2 di due componenti. La terza percentuale è $c_3 = 100 - c_1 - c_2$. Per rappresentare graficamente lo stato occorre combinare più diagrammi oppure considerarlo come un punto di uno spazio a quattro coordinate T, p, c_1, c_2. Questo metodo viene realmente usato in chimica. Le applicazioni della geometria multidimensionale alla chimica furono sviluppate dall'americano Gibbs e dalla scuola sovietica di chimica fisica di Kurnakov. L'introduzione di uno spazio multidimensionale risponde qui al desiderio di conservare le analogie e i ragionamenti geometrici resi possibili dai grafici.

Un esempio geometrico: una sfera è individuata da quattro dati, le tre coordinate del centro e il raggio.[21] La particolare geometria delle sfere, elaborata da vari matematici circa un secolo prima della stesura del libro, può quindi essere considerata una geometria quadridimensionale.

Il fondamento concreto del concetto è ormai chiaro. Se una figura, lo stato di un sistema o un altro oggetto è individuato da n dati indipendenti, lo si può interpretare come un punto di uno spazio a n dimensioni. Il vantaggio è simile a quello dei grafici ordinari: permette di applicare analogie e metodi geometrici allo studio dei fenomeni.

Non vi è dunque nulla di mistico nel concetto matematico di spazio multidimensionale. È un'astrazione sviluppata per descrivere geometricamente ciò che non ammette una rappresentazione semplice nello spazio ordinario. Ha un fondamento reale ed è nata dalle esigenze della scienza, non da un gioco ozioso dell'immaginazione. Sfere e miscele di tre gas, per esempio, richiedono più dati per essere descritte; la famiglia di tutti questi oggetti è quindi multidimensionale. Come un punto che si muove nello spazio cambia le proprie tre coordinate, così una sfera che si sposta, si dilata o si contrae cambia le quattro quantità che la determinano.

Nei paragrafi seguenti torneremo sulla geometria multidimensionale. Qui importa comprenderla come metodo di descrizione matematica di oggetti e fenomeni reali. L'idea, sfruttata da certi letterati e spiritisti, di uno spazio quadridimensionale nel quale sarebbe immerso il nostro spazio reale non ha rapporto con questo concetto matematico. Se vi si può vedere un legame con la scienza, è soltanto quello di una deformazione fantastica dei concetti scientifici.

5. Poliedri nello spazio a *n* dimensioni.

La geometria multidimensionale nacque come estensione formale della geometria analitica a un numero arbitrario di variabili. Ma questo non bastava: si voleva generalizzare soprattutto il metodo geometrico di ricerca, non soltanto i concetti. Occorreva quindi un'esposizione puramente geometrica, indipendente dall'apparato analitico. La diede per primo lo svizzero Schläfli nel 1852, studiando fra l'altro i poliedri regolari multidimensionali. Il suo lavoro non fu apprezzato dai contemporanei, perché richiedeva una concezione sufficientemente astratta della geometria. Gli sviluppi successivi chiarirono pienamente il rapporto fra approccio analitico e geometrico. Senza approfondire la questione, presenteremo alcuni esempi di trattazione geometrica.

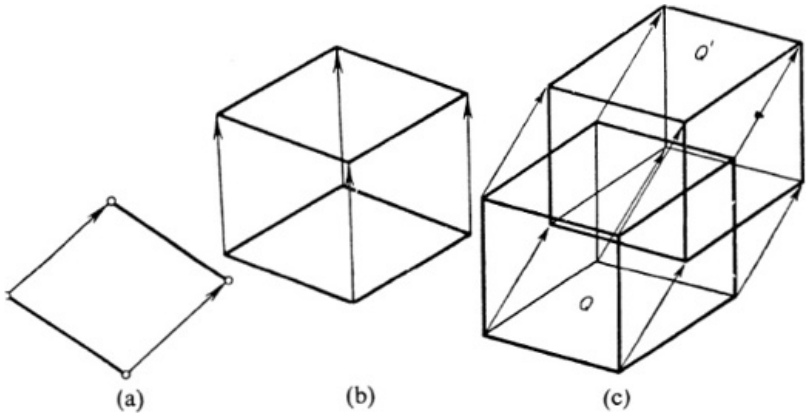


Figura 23.

Consideriamo la costruzione di un cubo a *n* dimensioni. Traslando un segmento nel piano, perpendicolarmente a se stesso, di una distanza uguale alla sua lunghezza, si descrive un quadrato, cioè un cubo bidimensionale (figura 23a). Traslando il quadrato perpendicolarmente al proprio piano di una distanza uguale al lato, si descrive un cubo tridimensionale (figura 23b). Per ottenere un cubo quadridimensionale, prendiamo un cubo

in un piano tridimensionale dello spazio a quattro dimensioni e trasliamolo perpendicolarmente a quel piano di una distanza uguale allo spigolo. Una retta è perpendicolare a un piano di dimensione *k* se è perpendicolare a tutte le rette del piano passanti per il piede. La figura 23c rappresenta simbolicamente la costruzione: *Q* e *Q′* sono il cubo iniziale e quello finale; i segmenti che uniscono i vertici corrispondenti sono i percorsi dei vertici durante la traslazione. Il cubo quadridimensionale ha 16 vertici, otto in *Q* e otto in *Q′*; 32 spigoli, dodici in ciascuno dei due cubi e otto laterali; otto facce tridimensionali cubiche, le sei generate dalle facce del cubo in movimento più quella iniziale e quella finale. Ha inoltre 24 facce quadrate: sei in *Q*, sei in *Q′* e dodici descritte dagli spigoli durante il movimento.

Il cubo quadridimensionale ha dunque otto facce tridimensionali, 24 bidimensionali, 32 unidimensionali, cioè spigoli, e 16 vertici. Ogni faccia è un «cubo» della dimensione corrispondente: cubo ordinario, quadrato, segmento o punto, considerato come cubo di dimensione zero.

Traslando analogamente un cubo quadridimensionale «nella quinta dimensione», si ottiene un cubo a cinque dimensioni. Ripetendo la costruzione si ottengono cubi di qualsiasi dimensione. Tutte le facce di un cubo a *n* dimensioni sono cubi di dimensione inferiore, fino agli spigoli e ai vertici. Il lettore può determinare il numero delle facce di ciascuna dimensione: è facile vedere che quelle di dimensione $n - 1$ sono $2n$ e che i vertici sono 2^n. Quanti sono, per esempio, gli spigoli?

Consideriamo un altro poliedro. Nel piano, il poligono con il minor numero di vertici è il triangolo. Per ottenere il poliedro tridimensionale più semplice basta prendere un punto esterno al suo piano e unirlo a tutti i punti del triangolo: i segmenti riempiono una piramide triangolare, cioè un tetraedro (figura 24). In quattro dimensioni prendiamo un tetraedro *T* in un piano tridimensionale e un punto *O* esterno a quel piano; uniamo *O* a tutti i punti di *T*. La costruzione è rappresentata simbolicamente a destra nella figura 24. Ciascun segmento non può avere un secondo punto comune con *T*, altrimenti giacerebbe interamente nel piano tridimensionale che lo contiene. I segmenti «entrano nella quarta dimensione» e formano il più semplice poliedro quadridimensionale: il simplesso quadridimensionale. Esso ha cinque facce tridimensionali tetraedriche, una di base e quattro laterali; dieci facce triangolari, quattro nella base e sei laterali; dieci spigoli e cinque vertici.

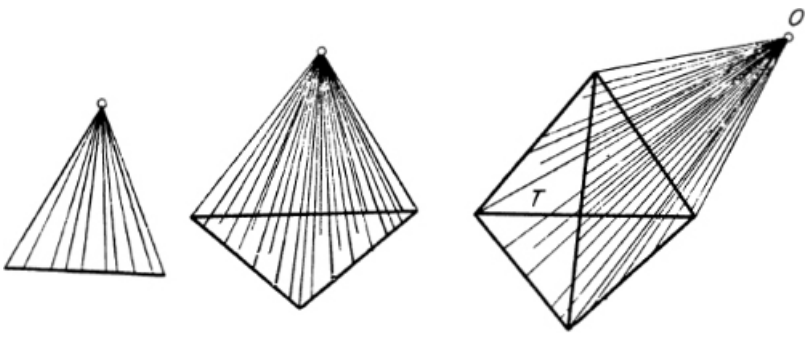


Figura 24.

Ripetendo la costruzione si ottiene il simplesso a n dimensioni, il più semplice poliedro di quella dimensione. Ha $n + 1$ vertici e tutte le sue facce sono simplessi di dimensione inferiore.[22]

È altrettanto facile generalizzare prismi e piramidi. Traslando un poligono piano nella terza dimensione si descrive un prisma; traslando un poliedro tridimensionale nella quarta si descrive un prisma quadridimensionale, rappresentato simbolicamente nella figura 25. Il cubo quadridimensionale ne è un caso particolare.

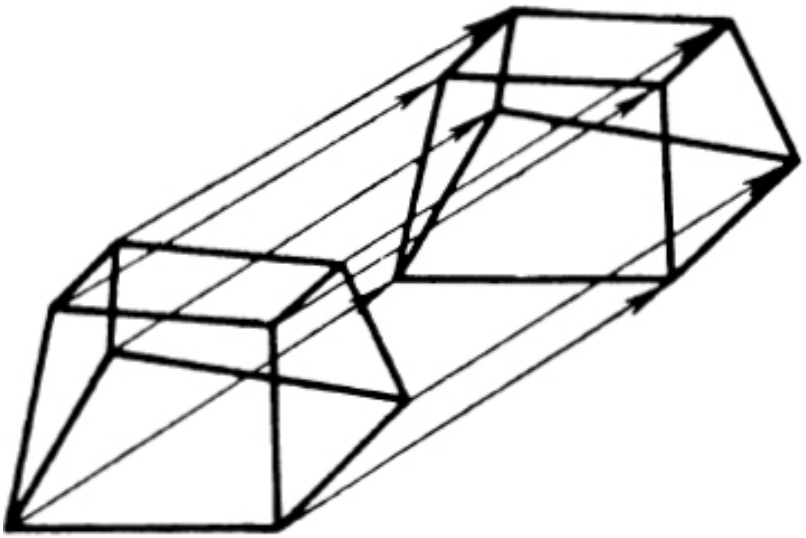


Figura 25.

Per costruire una piramide si prende un poligono Q e un punto O esterno al suo piano e si unisce O a tutti i punti di Q: i segmenti riempiono una piramide di base Q (figura 26). Allo stesso modo, nello spazio a quattro dimensioni, un poliedro tridimensionale Q e un punto O esterno al suo piano determinano una piramide quadridimensionale. Il simplesso quadridimensionale è una piramide con base tetraedrica.

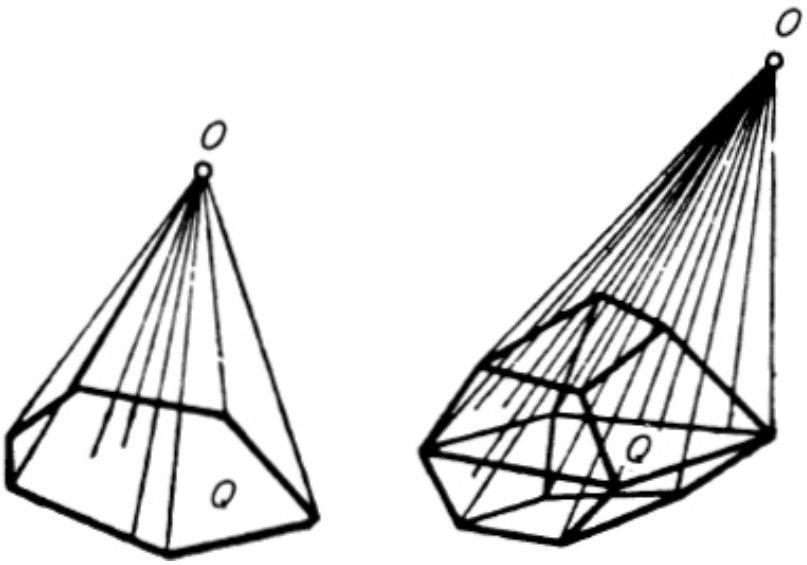


Figura 26.

Partendo da un poliedro Q di dimensione $n - 1$, si definiscono nello stesso modo un prisma e una piramide di dimensione n.

In generale, un poliedro a n dimensioni è una regione dello spazio delimitata da un numero finito di porzioni di iperpiani di dimensione $n - 1$. Un poliedro di dimensione k è analogamente una regione di un piano di dimensione k, delimitata da porzioni di piani di dimensione $k - 1$. Le sue facce sono a loro volta poliedri di dimensione inferiore.

La teoria dei poliedri multidimensionali generalizza quella dei poliedri ordinari ed è ricca di risultati concreti. Molti teoremi si estendono senza particolari difficoltà, mentre altri problemi diventano enormemente più difficili. Ricordiamo le profonde ricerche di G. F. Voronoj (1868-1908), nate anche da questioni di teoria dei numeri e proseguite dai geometri sovietici. Uno dei problemi, detto «problema di Voronoj», non era ancora completamente risolto all'epoca del libro.[23]

I poliedri regolari convessi mostrano quanto possano differire le dimensioni. Nel piano esistono poligoni regolari con un numero arbitrario di lati: le forme sono quindi infinite. In tre dimensioni esistono soltanto cinque poliedri regolari: tetraedro, cubo, ottaedro, dodecaedro e icosaedro. In quattro dimensioni le forme sono sei; in ogni dimensione maggiore sono soltanto tre: il simplesso regolare, con tutti gli spigoli uguali; il cubo; e l'analogo dell'ottaedro, ottenuto prendendo come vertici i centri delle facce di dimensione massima del cubo e formando il loro inviluppo convesso. La figura 27 mostra la costruzione in tre dimensioni. Per i poliedri regolari, dunque, le dimensioni due, tre e quattro hanno un carattere speciale.

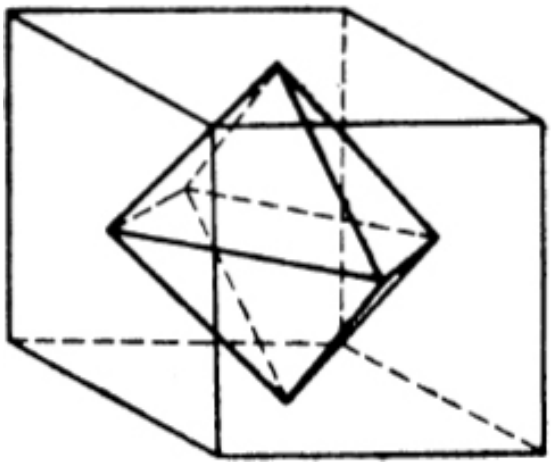


Figura 27.

6. Calcolo dei volumi.

Il volume di un corpo a n dimensioni si definisce in modo analogo al caso ordinario. È una grandezza numerica invariante per movimenti rigidi e additiva: se un corpo si compone di due parti senza sovrapposizione dei loro interni, il suo volume è la somma dei volumi delle parti. L'unità è il cubo di spigolo uno. Si dimostra poi che il cubo di spigolo a ha volume a^n, suddividendolo in strati di cubetti, come nel piano e nello spazio tridimensionale (figura 28). Poiché i cubetti si dispongono lungo n direzioni, il risultato è a^n.

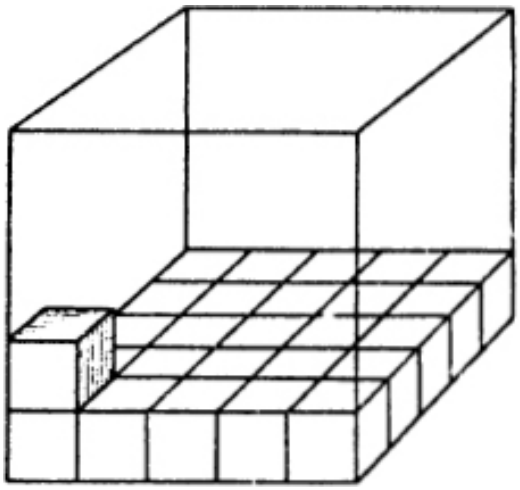


Figura 28.

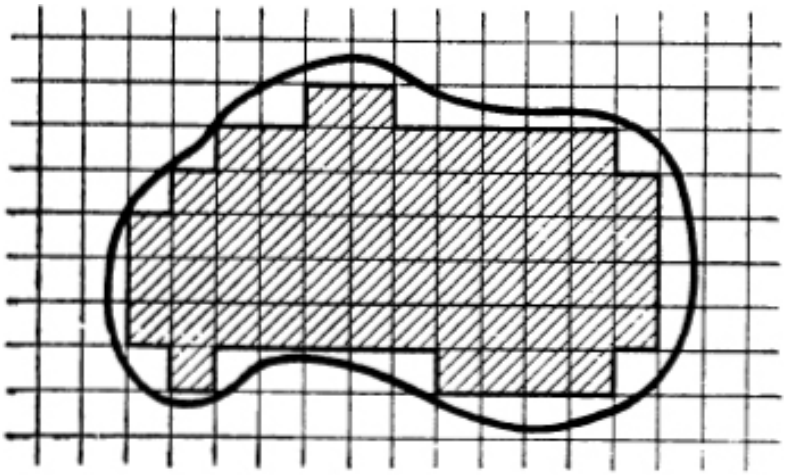


Figura 29.

Per definire il volume di un corpo più generale, lo si approssima con un'unione di molti cubetti a n dimensioni, come nella figura 29 una regione piana è approssimata da quadrati. Per i corpi misurabili nel senso qui considerato, il volume è il limite comune delle approssimazioni interne ed esterne quando la dimensione dei cubetti tende a zero.

Nello stesso modo si definisce il volume a k dimensioni di una figura contenuta in un piano di dimensione k. Dalla definizione segue una proprietà importante: se tutte le dimensioni lineari vengono moltiplicate per un fattore positivo λ, il volume a k dimensioni viene moltiplicato per λ^k.

Dividendo un corpo in strati paralleli, il volume totale è la somma dei volumi degli strati:

$$V = \Sigma V_i .$$

Il volume di ciascuno strato è approssimativamente il prodotto del suo spessore Δh_i per il volume a $n - 1$ dimensioni, cioè l'«area», della sezione corrispondente S_i. Quindi il volume complessivo è approssimato dalla somma

$$V \approx \sum S_i \, \Delta h_i \, .$$

Passando al limite quando tutti gli spessori Δh_i tendono a zero, otteniamo

$$V = \int_0^H S(h)\, dh, \qquad (9)$$

dove H è l'estensione del corpo nella direzione perpendicolare alle sezioni.

Il procedimento è identico a quello tridimensionale. In un prisma tutte le sezioni sono congruenti e la loro «area» non dipende da h: quindi $V = SH$, prodotto dell'«area» della base per l'altezza. Per una piramide a n dimensioni di altezza H e base di «area» S, tagliamo con un piano parallelo alla base a distanza h dal vertice. Si ottiene una piramide più piccola, di altezza h, con base di «area» $s(h)$. Essa è simile all'originale con rapporto h/H. Il volume a $n - 1$ dimensioni della sua base è dunque

$$s(h) = \left(\frac{h}{H}\right)^{n-1} S,$$

poiché moltiplicare le dimensioni lineari di una figura di dimensione $n - 1$ per λ ne moltiplica il volume per λ^{n-1}.

Per la (9), il volume della piramide è

$$V = \int_0^H s(h)\, dh,$$

e quindi

$$V = \int_0^H S\left(\frac{h}{H}\right)^{n-1} dh = \frac{S}{H^{n-1}} \int_0^H h^{n-1}\, dh = \frac{S}{H^{n-1}} \cdot \frac{H^n}{n} = \frac{1}{n} SH;$$

Il volume di una piramide a n dimensioni è dunque $1/n$ del prodotto dell'altezza per il volume a $n - 1$ dimensioni della base. Per $n = 2$ e $n = 3$ ritroviamo le formule note: l'area del triangolo è metà del prodotto di base e altezza; il volume della piramide ordinaria è un terzo del prodotto dell'area di base per l'altezza.

Una sfera piena può essere approssimata mediante piramidi sottilissime con vertice comune nel centro. Le altezze tendono al raggio R e le aree σ_i delle basi ricoprono approssimativamente la superficie sferica totale S. Poiché ciascuna piramide ha volume approssimativamente $R\sigma_i/n$, sommando troviamo

$$V \approx \frac{1}{n} R \sum \sigma_i \approx \frac{1}{n} RS.$$

Al limite otteniamo la formula esatta $V = RS/n$: il volume della sfera piena è $1/n$ del prodotto del raggio per l'area della superficie. Per $n = 2$ o 3 la relazione è ben nota.[24]

Ricordiamo una proprietà importante, dimostrabile anche in dimensione n con metodi analoghi a quelli tridimensionali: fra i corpi di volume assegnato, la sfera ha area superficiale minima; nelle usuali condizioni di regolarità, l'uguaglianza caratterizza la sfera.

7. Geometria «superiore» a *n* dimensioni.

Finora abbiamo considerato la geometria elementare, ma si può sviluppare anche una teoria generale delle curve e delle superfici. Nello spazio a n dimensioni, queste possono avere dimensioni da uno, nel caso delle curve, fino a $n - 1$. Una curva parametrizzata si definisce mediante coordinate dipendenti continuamente da un parametro t:

$$x_1 = x_1(t), \quad x_2 = x_2(t), \cdots, \quad x_n = x_n(t).$$

La curva è, per così dire, la traiettoria di un punto nello spazio a n dimensioni al variare di t. Se lo spazio rappresenta gli stati di un sistema fisico, come negli esempi del punto 4, la curva rappresenta una successione continua di stati o l'evoluzione del sistema rispetto al parametro, per esempio il tempo. È la generalizzazione dei grafici ordinari.

Per una curva sufficientemente regolare e non degenere, a ogni punto si associano non soltanto la tangente, ma anche i piani osculatori di dimensione da 2 a $n - 1$. Il modo in cui questi piani ruotano al crescere della lunghezza d'arco definisce le curvature corrispondenti. Si hanno così $n - 1$ piani, dalla tangente al piano osculatore di dimensione $n - 1$, e $n - 1$ curvature generalizzate. La geometria differenziale multidimensionale è quindi molto più ricca di quella tridimensionale.

Fin qui abbiamo parlato della generalizzazione diretta della geometria euclidea. Ma anche la geometria di Lobačevskij, quella proiettiva e le altre geometrie si estendono a un numero arbitrario di dimensioni.

§8. Ampliamento del campo della geometria

1. Lo spazio dei colori.

Discutendo il significato concreto degli spazi a *n* dimensioni abbiamo incontrato il problema dell'ampliamento della geometria e del concetto generale di spazio. Prima di darne una definizione, consideriamo alcuni esempi.

L'esperienza mostra, come aveva già osservato M. V. Lomonosov, che la visione umana normale è tricromatica: una percezione cromatica C si descrive mediante tre componenti fondamentali, indicate qui come rosso R, verde G e blu B, con determinate intensità.[25] Indicandole con x, y, z, scriviamo $C = xR + yG + zB$. Come un punto può spostarsi in tre direzioni indipendenti, così una percezione cromatica può variare modificando le tre componenti. Si parla quindi, per analogia, di «spazio tridimensionale dei colori», nel quale le intensità svolgono il ruolo di coordinate. Nel modello semplificato di miscelazione additiva con tre luci primarie fissate, le intensità sono non negative e descrivono i colori ottenibili da quelle primarie; non necessariamente tutti i colori percepibili. Per $x = y = z = 0$ si ha il nero, corrispondente all'assenza di luce.

Una variazione continua del colore si rappresenta come una linea in questo spazio. I colori dell'arcobaleno ne formano un esempio. Anche le percezioni prodotte da un oggetto uniformemente colorato, variando continuamente l'intensità dell'illuminazione a composizione spettrale costante, formano una linea: cambia l'intensità della percezione, mentre resta invariata la cromaticità.

Dati due colori, per esempio rosso R e bianco W, mescolandoli in proporzioni variabili[26] si ottiene una successione continua dal rosso al bianco, che possiamo chiamare segmento RW. Dire che il rosa si trova fra il rosso e il bianco ha un significato chiaro.

Nascono così figure e relazioni geometriche elementari nello spazio dei colori. Un «punto» è un colore; il «segmento» AB è l'insieme delle miscele di A e B; dire che D appartiene al segmento significa che è una loro miscela. La combinazione di tre colori dà una porzione di piano, un «triangolo cromatico». Mediante le coordinate x, y, z, le linee e i piani cromatici si descrivono con formule analoghe a quelle della geometria analitica ordinaria.[27]

Nel modello valgono le relazioni euclidee sulla disposizione di punti e segmenti: esse costituiscono una geometria affine. Possiamo quindi dire che l'insieme delle percezioni cromatiche considerate realizza una parte di uno spazio affine. La precisazione è necessaria perché le intensità delle primarie fissate sono non negative: il modello corrisponde alla regione in cui tutte e tre le coordinate sono positive o nulle.

Abbiamo inoltre un'idea naturale della differenza fra colori: il rosa pallido è più vicino al bianco del rosa intenso, il cremisi è più vicino al rosso che al blu. Questa nozione qualitativa di distanza può diventare quantitativa. Definirla però mediante la formula euclidea $r = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2}$ non rispecchia in generale la percezione: due

colori a uguale distanza secondo la formula possono differire percettivamente dal colore di riferimento in misura diversa. La distanza deve riflettere le relazioni reali fra le percezioni.

Seguendo questo principio si introduce una misura particolare. Quando un colore cambia continuamente, la variazione non viene percepita subito, ma soltanto dopo avere superato una soglia di discriminazione. Si suppone allora che i colori posti esattamente alla soglia rispetto a un colore dato siano tutti alla stessa distanza da esso. Si è così condotti a misurare la distanza mediante il minor numero di soglie successive che separano i due colori. La lunghezza di una linea cromatica si misura contando tali piccoli passi e la distanza fra due colori come lunghezza minima dei percorsi che li uniscono, analogamente alla distanza geometrica.

Le lunghezze e le distanze vengono quindi misurate mediante passi molto piccoli, idealizzati come infinitesimi.

Si definisce così una particolare geometria non euclidea dello spazio dei colori. Il suo significato è concreto: descrive in linguaggio geometrico proprietà della risposta visiva agli stimoli luminosi.

Il concetto di spazio dei colori era sorto circa un secolo prima della stesura del libro. Ne studiarono la geometria molti fisici, fra i quali Helmholtz e Maxwell. Queste ricerche proseguono e hanno interesse sia teorico sia pratico: forniscono basi matematiche per distinguere segnali cromatici, trattare i colori nell'industria tessile e affrontare altri problemi.

2. Spazi delle fasi in fisica e chimica.

Passiamo a un altro esempio già accennato nel paragrafo precedente.

Supponiamo di studiare un sistema fisico-chimico, come una miscela di gas o una lega, il cui stato sia determinato da n grandezze indipendenti: per esempio pressione, temperatura e concentrazioni dei componenti. Si dice che il sistema ha n gradi di libertà, perché il suo stato può cambiare in n direzioni indipendenti, variando ciascuna di queste grandezze. Esse svolgono il ruolo di coordinate; l'insieme degli stati può quindi essere considerato uno spazio a n dimensioni, detto qui spazio delle fasi del sistema.

I cambiamenti continui di stato, cioè i processi del sistema, sono rappresentati da linee in questo spazio. Gli insiemi di stati caratterizzati da particolari proprietà ne costituiscono regioni; gli stati di confine fra due regioni formano una superficie.

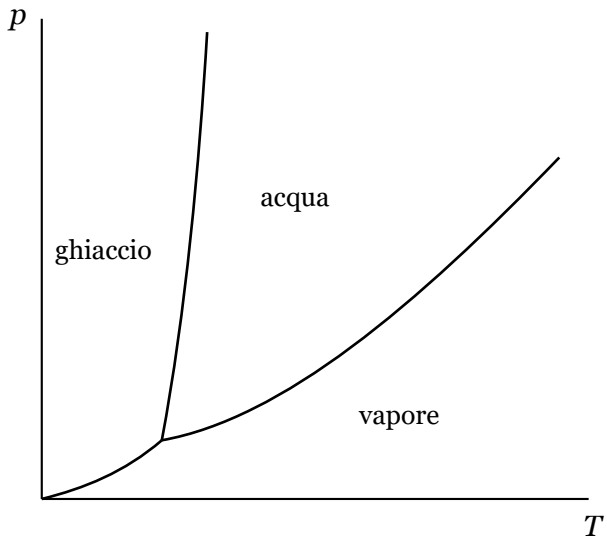


Figura 30.

In chimica fisica è importante studiare la forma e i rapporti di adiacenza delle regioni corrispondenti a stati qualitativamente diversi. Le superfici che le separano rappresentano transizioni come fusione, evaporazione o precipitazione. Con due gradi di libertà, lo stato si rappresenta con un punto del piano. Per una sostanza omogenea, per esempio, possiamo usare pressione p e temperatura T come coordinate. Studiamo allora le curve che separano le regioni delle diverse fasi. Per l'acqua esse corrispondono a ghiaccio, liquido e vapore (figura 30); le curve di separazione rappresentano fusione e solidificazione, evaporazione e condensazione, sublimazione e deposizione del ghiaccio.

Per studiare sistemi con molti gradi di libertà occorrono i metodi della geometria multidimensionale.

Il concetto di spazio delle fasi si applica anche ai sistemi meccanici e, più generalmente, a ogni sistema i cui stati possibili formino una famiglia continua. Nella teoria cinetica dei gas si considerano, per esempio, sistemi di particelle materiali, le molecole. Lo stato istantaneo di una particella è determinato da posizione e velocità: tre coordinate e tre componenti della velocità, in tutto sei grandezze. Lo stato di N particelle è descritto da $6N$ grandezze. Poiché le molecole sono numerosissime, $6N$ è enorme; ciò non impedisce ai fisici di parlare dello spazio delle fasi a $6N$ dimensioni del sistema.

Un punto di questo spazio descrive l'intero sistema di molecole, con tutte le posizioni e velocità; il movimento del punto ne descrive l'evoluzione. Questa rappresentazione astratta è molto utile in numerosi sviluppi teorici. Lo spazio delle fasi ha quindi un posto stabile negli strumenti delle scienze esatte e si applica a problemi diversi.

3. Generalizzazione della geometria.

Gli esempi ci permettono ora di comprendere come si amplia il campo della geometria.

Supponiamo di voler studiare una famiglia continua di oggetti, eventi o stati: per esempio i colori o gli stati di un insieme di molecole. Le relazioni fra essi possono essere analoghe alle relazioni spaziali ordinarie, come la distanza fra colori o la disposizione reciproca delle regioni di uno spazio delle fasi. Astraendo dalle particolarità qualitative degli oggetti e considerando soltanto queste relazioni, possiamo interpretare la famiglia come uno spazio. I suoi «punti» sono gli oggetti, eventi o stati stessi; una «figura» è un insieme di tali punti, come la linea dei colori dell'arcobaleno o la regione del vapore nello spazio degli stati dell'acqua. La geometria è determinata dalle relazioni analoghe a quelle spaziali: nello spazio dei colori, per esempio, dalle leggi di miscelazione e dalle distanze percettive.

Il valore concreto di questa prospettiva consiste nel poter applicare concetti e metodi geometrici a fenomeni molto diversi, ampliandone enormemente il campo d'impiego. Il termine «spazio» assume così due significati scientifici: da un lato lo spazio reale ordinario, che l'autore considera forma universale di esistenza della materia; dall'altro uno spazio astratto, una famiglia di oggetti dello stesso genere fra i quali sono definite relazioni di tipo spaziale.

Anche lo spazio ordinario, in una rappresentazione semplificata, può essere visto come una famiglia di stati: tutte le posizioni possibili di un corpo infinitamente piccolo, un «punto materiale». Non intendiamo darne così una definizione, ma chiarire il legame fra i due significati. Approfondiremo subito il concetto di spazio astratto; nell'ultima parte del capitolo discuteremo il rapporto fra geometria astratta e spazio reale.

4. Spazi generalizzati in matematica.

Le applicazioni più ampie degli spazi astratti si trovano nella matematica stessa. In geometria si considerano spazi di figure: lo spazio delle sfere già ricordato, quello delle rette e così via.

Il metodo è particolarmente fecondo nella teoria dei poliedri. Nel capitolo VII, §5, abbiamo ricordato il teorema di esistenza di un poliedro convesso con uno sviluppo assegnato. La dimostrazione considera due spazi: quello dei poliedri e quello degli sviluppi. I poliedri convessi con un dato numero di vertici sono trattati come punti di uno spazio; analogamente gli sviluppi ammissibili sono punti di un altro. La costruzione del poliedro a partire dallo sviluppo stabilisce una corrispondenza fra i due spazi. Occorre mostrare che a ogni sviluppo ammissibile corrisponde un poliedro, cioè a ogni punto del primo spazio un punto del secondo. Questo si dimostra mediante la topologia.

Molti altri teoremi sui poliedri si dimostrano in modo analogo. In diversi casi, fra cui quello appena citato, il metodo degli spazi astratti è il più semplice fra quelli noti. Rimane tuttavia abbastanza complesso da non poterlo esporre qui con maggiore precisione.

Il concetto generalizzato di spazio trova ampie applicazioni anche in analisi, algebra e teoria dei numeri. Il punto di partenza è la rappresentazione delle funzioni mediante curve. I valori

di una variabile si rappresentano su una retta; quelli di due variabili nel piano; quelli di *n* variabili nello spazio a *n* dimensioni. Alla n-upla (x_1, …, x_n) corrisponde il punto con quelle coordinate. Parliamo così di dominio delle variabili, insieme dei valori di una funzione, punti, linee o superfici di discontinuità. Non è soltanto un linguaggio: la rappresentazione geometrica rende intuitivi molti fatti analitici per analogia con lo spazio ordinario e permette di usare metodi geometrici di dimostrazione generalizzati a *n* dimensioni.

Lo stesso accade in algebra con equazioni in *n* incognite e funzioni algebriche di *n* variabili. Un'equazione lineare definisce un iperpiano; *m* equazioni definiscono *m* iperpiani, e ogni soluzione comune è un punto della loro intersezione. Questa può essere vuota, un punto, una retta o un piano di dimensione maggiore. Il problema della risolubilità di un sistema lineare diventa così un problema d'intersezione. L'approccio presenta numerosi vantaggi: l'algebra lineare, che studia equazioni e trasformazioni lineari, viene spesso esposta in larga parte geometricamente, come nel capitolo XVI.

5. Spazi a infinite dimensioni; definizione di «spazio».

Negli esempi precedenti abbiamo trattato come spazio una famiglia continua di colori, stati fisici, figure o valori di variabili. Ogni oggetto era determinato da un numero finito di dati indipendenti, e lo spazio aveva un numero finito di dimensioni.

All'inizio del Novecento si cominciarono a studiare anche spazi a infinite dimensioni: famiglie i cui elementi non sono descritti da un numero finito di parametri nel senso strutturale considerato. Il caso principale è quello degli spazi di funzioni.

Considerare una famiglia di funzioni come uno spazio è una delle idee fondamentali dell'analisi funzionale. Essa è estremamente feconda nella soluzione di molti problemi; il capitolo XIX sarà dedicato specificamente a questa disciplina.

Si possono considerare spazi di funzioni continue di una o più variabili, ma anche classi di funzioni discontinue. A seconda del problema si definisce una distanza in modi diversi. Il numero degli spazi funzionali possibili è illimitato e molti di essi sono effettivamente utilizzati.

Analogamente si considerano spazi di curve, di corpi convessi o dei moti possibili di un sistema meccanico. Nel capitolo VII, §5, abbiamo ricordato, nel contesto delle superfici convesse chiuse, teoremi sull'esistenza di almeno tre geodetiche chiuse e di infinite geodetiche fra due punti. Le dimostrazioni utilizzano spazi di curve sulla superficie: curve chiuse nel primo caso, curve con estremi assegnati nel secondo. Introducendo una distanza fra le curve, la loro famiglia diventa uno spazio al quale si applicano risultati profondi della topologia.

Formuliamo ora una conclusione generale.

In matematica chiamiamo «spazio», in senso ampio, una famiglia di oggetti dello stesso genere, come eventi, stati, funzioni, figure o valori di variabili, fra i quali sono definite relazioni analoghe a quelle spaziali: continuità, distanza e così via. Nel trattarla come spazio, astrarre significa trascurare tutte le proprietà degli oggetti salvo quelle determinate da queste

relazioni. Esse ne costituiscono la struttura, o «geometria». Gli oggetti sono i punti; gli insiemi di punti sono le figure.

La geometria di uno spazio astratto studia le proprietà determinate dalle relazioni prescelte. In uno spazio di funzioni continue, per esempio, si prescinde dalla struttura interna di ciascuna funzione: essa svolge il ruolo di un punto e in questo senso «non ha parti». Più precisamente, si trascurano le proprietà che non vengono considerate nella struttura dello spazio. Si studiano invece le relazioni fra funzioni, le loro distanze e le proprietà che da queste discendono.

Alla varietà degli oggetti e delle loro relazioni corrisponde un'illimitata varietà di spazi. Possiamo classificarli secondo le relazioni poste alla base della definizione. Senza tentare un elenco completo, presentiamo due classi fondamentali: spazi topologici e spazi metrici.

6. Spazi topologici.

Uno spazio topologico (si veda il capitolo XVIII) è un insieme di elementi, chiamati punti, nel quale è definita una relazione di aderenza di un punto a un insieme e, di conseguenza, fra insiemi. Si generalizza così la nozione intuitiva di vicinanza o aderenza delle figure nello spazio ordinario.

Già Lobačevskij, con notevole lungimiranza, indicava la vicinanza come la relazione geometrica più fondamentale: «La contiguità costituisce una proprietà caratteristica dei corpi e dà loro la stessa forma geometrica quando conserviamo questa proprietà senza considerare tutte le altre, essenziali o accidentali».[28] Per esempio, ogni punto della circonferenza è aderente all'interno del cerchio; due parti di un corpo connesso non sono separate nel senso topologico. Lo sviluppo successivo ha mostrato che l'aderenza è alla base di tutte le proprietà topologiche.

L'aderenza esprime l'idea che un punto sia arbitrariamente vicino a un insieme. Una famiglia di oggetti dotata di una nozione appropriata di continuità e vicinanza può dunque essere strutturata come spazio topologico.

Il concetto è estremamente generale: la topologia astratta rappresenta lo studio matematico più generale della continuità.

Una descrizione assiomatica può essere impostata nel modo seguente.

In un insieme R di elementi arbitrari, chiamati punti, si associano a ogni sottoinsieme M i suoi punti aderenti, richiedendo anzitutto le condizioni seguenti:

1. Ogni punto di M è aderente a M. È naturale che un insieme aderisca a ciascuno dei propri punti.

2. Se M_1 contiene M_2, ogni punto aderente a M_2 è aderente anche a M_1. In termini meno precisi: ingrandire un insieme non può ridurre l'insieme dei suoi punti aderenti.

A queste condizioni si aggiungono altri assiomi. Per ottenere precisamente una topologia nel senso usuale, l'insieme dei punti aderenti, detto chiusura, deve inoltre essere vuoto per l'insieme vuoto, rispettare l'unione di due insiemi e non cambiare applicando due volte la chiusura. Ulteriori condizioni distinguono poi varie classi di spazi topologici.

Mediante l'aderenza si definiscono molti concetti fondamentali della topologia e, più generalmente, della geometria. Le definizioni hanno un significato intuitivo chiaro. Eccone alcuni esempi.

1. Aderenza fra insiemi. Diciamo che due insiemi M_1 e M_2 aderiscono l'uno all'altro se almeno uno contiene un punto aderente all'altro. In questo senso, per esempio, una circonferenza aderisce all'interno del cerchio.

2. Connessione. Una figura, cioè un insieme M, è connessa se non può essere divisa in due parti non vuote che non aderiscono l'una all'altra. Un segmento è connesso; un segmento privato del punto medio non lo è.

3. Frontiera. La frontiera di M nello spazio R è l'insieme dei punti aderenti sia a M sia al suo complementare $R - M$. È la formulazione naturale dell'idea di confine.

4. Punto interno. Un punto di M è interno se non appartiene alla frontiera, cioè se non è aderente al complementare $R - M$.

5. Applicazione continua. Una trasformazione è continua se conserva l'aderenza: l'immagine di un punto aderente a un insieme è aderente all'immagine di quell'insieme. Questa esprime in forma precisa l'idea che una trasformazione continua non spezzi i rapporti di vicinanza.

Si potrebbero aggiungere altre definizioni importanti, come la convergenza di successioni di figure verso una figura o la dimensione di uno spazio.

I concetti geometrici più fondamentali si possono dunque definire mediante l'aderenza. Uno dei meriti della topologia consiste nel darne definizioni generali rigorose, rendendo possibile usare con precisione i ragionamenti suggeriti dall'intuizione della continuità.

La topologia studia le proprietà degli spazi, delle loro figure e delle trasformazioni che dipendono dalla relazione di aderenza.

La generalità di questa relazione fa della topologia una teoria capace di penetrare in molti rami della matematica, ovunque intervenga la continuità. Proprio per questo, nelle parti più astratte, essa oltrepassa i confini della geometria in senso stretto. Alla base resta però la generalizzazione di proprietà dello spazio reale; molti dei suoi risultati più potenti nascono da idee geometriche intuitive. Ne è un esempio il metodo di approssimare figure generali con poliedri, sviluppato da P. S. Aleksandrov ed esteso, in forma astratta, a classi molto generali di spazi topologici.

Quando uno studioso scopre nel proprio campo un modo naturale di definire vicinanza o aderenza, dispone subito dell'ampio apparato della topologia. Può così ottenere risultati tutt'altro che ovvi anche per il suo problema particolare.

7. Spazi metrici.

Uno spazio metrico è un insieme di elementi, chiamati punti, fra i quali è definita una distanza: a ogni coppia X, Y si associa un numero $r(X, Y)$ che soddisfa gli assiomi seguenti.

1. $r(X, Y) = 0$ se e solo se X e Y coincidono.

2. Per tre punti arbitrari X, Y, Z,

$$r(X, Y) + r(Y, Z) \geqslant r(Z, X).$$

Questa condizione è detta disuguaglianza triangolare, per analogia con la proprietà della distanza ordinaria fra tre punti A_1, A_2, A_3 dello spazio euclideo (figura 31):

$$r(A_1 , A_2) + r(A_2 , A_3) \geqslant r(A_3 , A_1).$$

Esempi di spazi metrici sono:

1. Lo spazio euclideo di dimensione arbitraria n.

2. Lo spazio di Lobačevskij.

3. Una superficie connessa con la propria distanza intrinseca, nelle condizioni descritte nel capitolo VII, §4.

4. Lo spazio C delle funzioni continue su un intervallo chiuso e limitato, con distanza $r(f_1, f_2) = \max |f_1(x) - f_2(x)|$.

5. Lo spazio di Hilbert, che descriveremo nel capitolo XIX: un analogo infinito-dimensionale dello spazio euclideo.

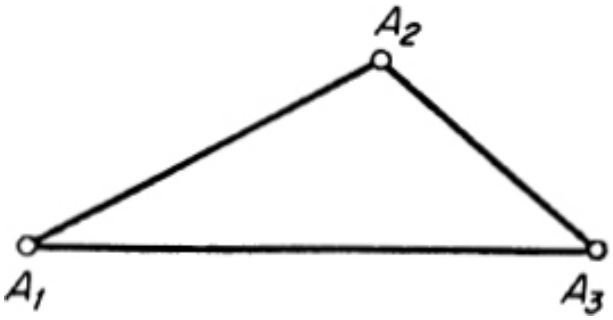


Figura 31.

Lo spazio di Hilbert è fra gli spazi più importanti dell'analisi funzionale. È legato alle serie di Fourier e, più generalmente, agli sviluppi in funzioni ortogonali: le coordinate x_1, x_2, ... sono allora i coefficienti dello sviluppo. Ha inoltre un ruolo fondamentale nella fisica matematica e nella meccanica quantistica. Gli stati puri di un sistema atomico, per esempio l'atomo di idrogeno, vengono rappresentati mediante vettori di uno spazio di Hilbert, con l'identificazione di vettori che differiscono soltanto per un fattore di fase.

Gli esempi si potrebbero moltiplicare. Nel prossimo paragrafo incontreremo una classe importante, gli spazi riemanniani; quelli già dati bastano però a mostrare l'ampiezza del concetto di spazio metrico.

In uno spazio metrico si definiscono sempre i concetti topologici e, inoltre, concetti propriamente metrici, come la lunghezza di una curva. Essa si definisce come nel caso ordinario: si prendono punti X_1, X_2, ..., X_n in ordine lungo la curva e si sommano le distanze fra punti consecutivi. La lunghezza è l'estremo superiore di tali somme, oppure il limite delle somme lungo suddivisioni opportunamente raffinate; può anche essere infinita. Si conservano così le proprietà fondamentali della lunghezza.

8. Vantaggi del metodo geometrico.

Oltre agli spazi topologici e metrici, si studiano molti altri tipi di spazi. Nel §6 abbiamo incontrato quelli dotati di un gruppo di trasformazioni, come gli spazi proiettivi e affini. In essi si può definire un'equivalenza fra figure: due figure sono «uguali» nel senso della teoria se una trasformazione del gruppo porta l'una nell'altra.

Non approfondiremo le numerose definizioni possibili; per esse rimandiamo alla letteratura sui vari rami della geometria contemporanea.

Ma perché ampliare tanto i concetti geometrici? Perché, per esempio, introdurre lo spazio delle funzioni continue? Non basterebbero i mezzi abituali dell'analisi, senza ricorrere a simili astrazioni?

La risposta generale è che introdurre uno spazio permette di applicare una grande varietà di concetti e metodi geometrici.

Questi metodi si fondano, in ultima analisi, su idee intuitive e ne conservano i vantaggi anche in forma astratta. Ciò che l'analista raggiunge con lunghi calcoli può talvolta essere colto subito dal geometra. Un grafico che rende evidente una dipendenza fra grandezze ne è l'esempio più elementare. Il metodo geometrico ha un carattere sintetico: permette di considerare insieme le relazioni del problema. Nelle teorie astratte l'intuizione immediata si attenua, ma rimangono il ragionamento per analogia e questa visione d'insieme.

Il lettore ha già incontrato immagini geometriche nell'analisi, la rappresentazione dei numeri complessi, il ragionamento topologico sul teorema fondamentale dell'algebra e altre applicazioni. In tutti questi esempi si riconoscono i vantaggi appena descritti. Ricordiamo anche gli esempi iniziali del §7 e quelli del punto 5 di questo paragrafo. Presentiamo ora un problema analitico risolvibile mediante uno spazio di funzioni.

La topologia dimostra che, se una regione piana chiusa ha la forma di un disco deformato e la si trasforma continuamente in se stessa, almeno un suo punto resta fisso. La trasformazione può anche sovrapporre parti della regione: non occorre che sia biunivoca. È un fatto puramente topologico.

Consideriamo ora un problema apparentemente lontano dalla geometria: trovare una funzione $y(x)$ che soddisfi l'equazione differenziale[29]

$$y' = f(x, y) \qquad (10)$$

e la condizione iniziale $y(0) = 0$.

In modo equivalente, possiamo cercare una soluzione dell'equazione integrale

$$y = \int_0^x f[t, y(t)]\, dt. \qquad (11)$$

Sorge naturalmente la domanda: esiste una funzione che soddisfi questa condizione?

Rappresentiamo ogni funzione continua $y(x)$ con un punto di uno spazio astratto. Il risultato del calcolo dell'integrale

$$\int_0^x f[t, y(t)]\, dt = z(x)$$

è ancora una funzione continua di x, quindi un punto dello stesso spazio. A funzioni y diverse corrispondono in generale funzioni z diverse: abbiamo definito una trasformazione dello spazio in se stesso. Trovare una soluzione della (11) equivale a trovare un punto che, dopo la trasformazione, coincida con la propria posizione iniziale, cioè un punto fisso.

Un problema naturale di equazioni differenziali si è trasformato in una questione relativa a uno spazio funzionale. L'analogia con il teorema precedente suggerisce di cercarne la soluzione mediante proprietà topologiche.

Opportune ricerche topologiche conducono così a dimostrazioni particolarmente brevi di molti teoremi di esistenza per equazioni differenziali. In particolare, se $f(x, y)$ è continua in un intorno del dato iniziale, la (10) ammette almeno una soluzione locale con quel dato iniziale.

§9. Geometria riemanniana

1. Storia della geometria riemanniana.

L'idea di considerare ogni famiglia continua di fenomeni omogenei come uno spazio fu espressa per la prima volta da Riemann nella lezione *Sulle ipotesi che stanno a fondamento della geometria*, tenuta a Gottinga nel 1854. Era una lezione di abilitazione, richiesta prima di assumere l'insegnamento. Senza calcoli né dimostrazioni, Riemann delineò una vasta teoria originale, oggi chiamata geometria riemanniana. Si racconta che soltanto l'anziano Gauss, fra gli ascoltatori, ne comprendesse la portata. Riemann elaborò l'apparato formale in un altro lavoro, applicandolo alla conduzione del calore: la teoria nacque quindi in stretto rapporto con la fisica matematica. Dopo il passo decisivo di Lobačevskij, le idee di Riemann segnarono la tappa successiva. Non furono però subito apprezzate: la lezione e i lavori sul calore vennero pubblicati postumi nel 1868. Nello stesso anno apparve il modello di Beltrami; nel 1870 quello di Klein. Nel 1872 Klein espose la concezione delle geometrie come studi delle proprietà invarianti rispetto a gruppi di trasformazioni, mentre la geometria multidimensionale si consolidava. Gli anni Settanta dell'Ottocento furono quindi un periodo decisivo: idee maturate nei cinquant'anni precedenti vennero finalmente comprese da una cerchia ampia di matematici ed entrarono stabilmente nella disciplina.

Il lavoro di Riemann fu proseguito e, alla fine dell'Ottocento, la geometria riemanniana era già molto sviluppata e applicata alla meccanica e alla fisica. Nel 1915 Einstein impiegò nella relatività generale una geometria pseudo-riemanniana per descrivere la gravitazione. L'evento attirò grande attenzione su queste teorie, stimolandone lo sviluppo e le generalizzazioni.

2. Le idee fondamentali.

Se si lasciano da parte i dettagli tecnici, le idee di Riemann sono semplici nella loro essenza. Questa semplicità interna caratterizza molte grandi idee: considerare una geometria fondata sulla negazione del quinto postulato, l'evoluzione degli organismi o la struttura atomica della materia. Sono idee semplici e insieme complesse: devono svilupparsi in un campo vasto senza essere costrette in schemi rigidi; fondarle, elaborarle e applicarle richiede inoltre enorme lavoro, ingegnosità e strumenti specialistici. Nella geometria riemanniana questi strumenti comprendono formule complesse, che qui eviteremo. Riemann parte da una famiglia continua di fenomeni, considerata come spazio. Le quantità che individuano un fenomeno ne sono le coordinate, come le intensità x, y, z nel modello del colore $C = xR + yG + zB$. Se le coordinate indipendenti sono n, si parla di spazio a n dimensioni. In esso si considerano linee e se ne misura la lunghezza mediante passi infinitesimi, analogamente alle curve ordinarie.

Per misurare così le lunghezze basta una regola che assegni, in ogni punto, la lunghezza di uno spostamento infinitesimo. Tale regola si chiama metrica. Il caso più semplice è quello in cui essa ha, nel punto, la stessa forma che nello spazio euclideo. Lo spazio è allora euclideo «nell'infinitamente piccolo»: in regioni abbastanza piccole le relazioni euclidee valgono

approssimativamente, con precisione crescente al diminuire della regione. Uno spazio dotato di una metrica di questo tipo si chiama riemanniano, e riemanniana è la sua geometria.

L'esempio più semplice è una superficie liscia con la propria geometria intrinseca, che costituisce una geometria riemanniana bidimensionale. Vicino a un punto, la superficie differisce poco dal piano tangente, tanto meno quanto più piccola è la regione considerata. Di conseguenza, anche la geometria locale differisce poco da quella piana. Su regioni estese, invece, può essere molto diversa, come mostrano sfera e pseudosfera e come abbiamo spiegato nel capitolo VII, §4. La geometria riemanniana generalizza questa situazione a un numero arbitrario di dimensioni. Anche uno spazio riemanniano tridimensionale può essere euclideo in prima approssimazione vicino a ogni punto e differire dallo spazio euclideo su grandi regioni. La lunghezza di una circonferenza può non essere proporzionale al raggio, salvo approssimativamente per piccoli raggi; la somma degli angoli di un triangolo può differire da due retti. I lati del triangolo sono qui geodetiche, localmente linee di minima lunghezza fra gli estremi.

Si può quindi ipotizzare che lo spazio fisico sia descritto con grande precisione dalla geometria euclidea soltanto in regioni piccole rispetto alle scale astronomiche. A scale maggiori può manifestarsi una geometria diversa. L'idea era già stata proposta da Lobačevskij; Riemann la estese a una classe molto più ampia di geometrie, nella quale quella di Lobačevskij compare come caso particolare.

La geometria riemanniana nacque dunque dalla sintesi di tre idee: la possibilità di geometrie diverse da quella euclidea, la geometria intrinseca delle superfici e gli spazi di dimensione arbitraria.

3. Misura delle distanze.

Per capire come si definisce matematicamente uno spazio riemanniano, ricordiamo la misura delle distanze nello spazio euclideo.

In un piano con coordinate cartesiane ortogonali x, y, il teorema di Pitagora esprime la distanza fra due punti le cui coordinate differiscono di Δx, Δy mediante

$$s = \sqrt{\Delta x^2 + \Delta y^2}.$$

Analogamente, nello spazio tridimensionale,

$$s = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}.$$

Nello spazio euclideo a n dimensioni vale la formula generale

$$s = \sqrt{\Delta x_1^2 + \cdots + \Delta x_n^2}\,.$$

È ora facile capire come definire la misura delle distanze in uno spazio riemanniano. La regola deve coincidere con quella euclidea soltanto al primo ordine, in un intorno infinitesimo di ciascun punto. Si giunge alla formulazione seguente.

Uno spazio riemanniano a n dimensioni è caratterizzato dal fatto che, vicino a ogni punto A, si possono introdurre coordinate $x_1, \ldots, x_n$ per le quali l'elemento di lunghezza in A ha la forma

$$ds = \sqrt{dx_1^2 + \cdots + dx_n^2}, \qquad (12)$$

dove $dx_1, \ldots, dx_n$ sono gli incrementi infinitesimi delle coordinate. In termini di punti vicini distinti, la distanza da A a X è data dalla formula euclidea con un errore relativo che tende a zero quando X tende ad A:

$$s(AX) = \sqrt{\Delta x_1^2 + \cdots + \Delta x_n^2} + \epsilon,$$

Qui ϵ è un termine piccolo rispetto al primo, nel senso che il loro rapporto tende a zero quando gli incrementi $\Delta x_1, \ldots, \Delta x_n$ tendono a zero.[30]

Questa è la caratterizzazione locale della metrica riemanniana, nel consueto quadro di coordinate differenziabili e coefficienti sufficientemente regolari. La differenza rispetto alla metrica euclidea è che la forma semplice vale nel singolo punto considerato; per ottenerla in punti diversi occorre in generale cambiare coordinate.[31] Preciseremo fra poco questa differenza.

L'approssimazione euclidea infinitesima permette di definire le grandezze geometriche come nella geometria intrinseca delle superfici, approssimando una piccola porzione con il piano tangente (capitolo VII, §4). Il volume infinitesimo si esprime come nel caso euclideo; il volume di una regione si ottiene integrandolo. La lunghezza di una curva si ottiene integrando l'elemento ds lungo la curva: è la formulazione analitica rigorosa dell'idea di riportarvi un regolo infinitesimo. L'angolo fra due curve in un punto comune è quello fra i rispettivi vettori tangenti, misurato con la metrica nel punto. Si possono inoltre considerare superfici regolari di dimensione da 2 a $n - 1$. Ognuna, con la metrica indotta, è a sua volta uno spazio riemanniano, come una superficie ordinaria lo è a due dimensioni.

Si può anche rappresentare uno spazio riemanniano come una superficie di dimensione n in uno spazio euclideo di dimensione sufficientemente alta, conservandone la geometria intrinseca. Il limite dimensionale $n(n + 1)/2$ citato qui dall'originale riguarda il teorema di immersione locale per metriche analitiche; non è una garanzia generale di rappresentazione globale per ogni metrica riemanniana.

4. La forma quadratica fondamentale.

Per esprimere analiticamente le grandezze geometriche occorre una formula della metrica valida in coordinate arbitrarie, senza cambiare sistema a ogni punto. La (12) assume infatti una forma semplice soltanto scegliendo coordinate adatte al punto considerato, il che è scomodo. Si può però dimostrare quanto segue.

Introduciamo coordinate y_1, ..., y_n in una regione dello spazio riemanniano. L'elemento di lunghezza dal punto A di coordinate y_1, ..., y_n al punto infinitamente vicino di coordinate $y_1 + dy_1$, ..., $y_n + dy_n$ è espresso dalla formula

$$ds = \sqrt{(\sum_{i,k=1...n} g_{ik}\, dy_i\, dy_k)}, \text{ ossia}$$
$$ds^2 = \sum g_{ik}\, dy_i\, dy_k. \quad (13)$$

dove i coefficienti g_{ik} sono funzioni delle coordinate del punto A.

Il secondo membro è una forma quadratica[32] nei differenziali delle coordinate. Per esteso si scrive:

$$\sum g_{ik}\, dy_i\, dy_k = g_{11}\, dy_1^2 + g_{12}\, dy_1\, dy_2 + g_{21}\, dy_2\, dy_1 + g_{22}\, dy_2^2 + \cdots .$$

Poiché $dy_1 dy_2 = dy_2 dy_1$, conviene scegliere $g_{12} = g_{21}$ e, in generale, $g_{ik} = g_{ki}$. È sempre possibile, perché conta soltanto la somma dei due termini, $(g_{ik} + g_{ki})dy_i dy_k$.

La forma è definita positiva: $ds^2 > 0$ ogni volta che i differenziali non sono tutti nulli.

Vale anche il viceversa. Se in uno spazio a n dimensioni l'elemento di lunghezza è dato dalla (13), con forma quadratica definita positiva e coefficienti regolari, lo spazio è riemanniano. Vicino a ogni punto A si possono cioè scegliere nuove coordinate x_1, ..., x_n in modo che, nel punto A, l'elemento di lunghezza abbia la forma semplice (12):

$$ds^2 = dx_1^2 + dx_2^2 + \cdots + dx_n^2 .$$

Una metrica riemanniana, euclidea nell'infinitamente piccolo, è dunque determinata da una forma quadratica definita positiva come la (13), con coefficienti g_{ik} dipendenti regolarmente dalle coordinate. Questo è il modo generale di assegnarla.

Una curva nello spazio è descritta facendo dipendere le n coordinate da un parametro t, variabile in un intervallo:

$$y_1 = y_1(t),\ y_2 = y_2(t),\ \cdots,\ y_n = y_n(t) \qquad (a \leqslant t \leqslant b). \qquad (14)$$

La sua lunghezza è espressa dall'integrale

$$s = \int ds = \int \sqrt{\sum g_{ik}\, dy_i\, dy_k} .$$

Per la curva definita dalle (14) abbiamo

$$dy_1 = y_1'\, dt, \cdots, dy_n = y_n'\, dt,$$

e quindi

$$s = \int_a^b \sqrt{\sum g_{ik}\, y_i' y_k'}\, dt. \qquad (15)$$

I coefficienti g_{ik} sono funzioni note delle coordinate, che dipendono a loro volta da t secondo le (14). L'integrando della (15) è quindi determinato dalla curva e il suo integrale ne fornisce la lunghezza.

La distanza fra due punti è l'estremo inferiore delle lunghezze delle curve che li uniscono. Quando il minimo è raggiunto, una curva minimizzante è una geodetica e svolge il ruolo del segmento rettilineo. In un intorno sufficientemente piccolo e opportunamente scelto, due punti sono uniti da un'unica geodetica minimizzante. La ricerca delle geodetiche conduce a un problema variazionale per l'integrale (15), del tipo studiato nel capitolo VIII. I metodi del calcolo delle variazioni forniscono le equazioni differenziali delle geodetiche e permettono di studiarne le proprietà generali.

Dimostriamo l'affermazione principale: in coordinate arbitrarie una metrica riemanniana è espressa dalla (13).

Siano y_1, ..., y_n coordinate in una regione. Fissiamo un punto A e scegliamo coordinate speciali x_1, ..., x_n per le quali l'elemento di lunghezza in A soddisfi la (12), ossia

$$ds^2 = dx_1^2 + dx_2^2 + \cdots + dx_n^2 . \qquad (16)$$

Le coordinate x_i si esprimono mediante le y_j con formule del tipo

$$\begin{aligned} x_1 &= f_1(y_1, y_2, \cdots, y_n), \\ x_2 &= f_2(y_1, y_2, \cdots, y_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= f_n(y_1, y_2, \cdots, y_n). \end{aligned}$$

Quindi

$$dx_1 = (\partial f_1/\partial y_1)\, dy_1 + (\partial f_1/\partial y_2)\, dy_2 + \ldots + (\partial f_1/\partial y_n)\, dy_n,$$

e analogamente per gli altri differenziali. Sostituendo nella (16), sviluppando i quadrati e raccogliendo i termini con dy_1^2, $dy_1\, dy_2$, dy_2^2, ecc., otteniamo

$$ds^2 = g_{11}\, dy_1^2 + 2g_{12}\, dy_1\, dy_2 + g_{22}\, dy_2^2 + \cdots + g_{nn}\, dy_n^2 ,$$

dove i coefficienti g_{11}, g_{12}, ..., g_{nn} dipendono dalle derivate parziali $\partial f_i/\partial y_j$. Non ci interessa qui la loro espressione esplicita. Questa è precisamente la (13) sviluppata: l'affermazione è dimostrata.

Mostriamo ora il viceversa: una forma (13) definita positiva può essere ridotta, in ogni punto, alla forma semplice (16) scegliendo opportunamente le coordinate. Sia

$$ds^2 = \sum g_{ik}\, dy_i\, dy_k ,$$

con g_{ik} funzioni delle coordinate e forma definita positiva. Fissato un punto, i coefficienti sono numeri assegnati e le variabili della forma sono i differenziali dy_1, ..., dy_n. L'algebra insegna che una forma quadratica definita positiva si riduce a somma di quadrati mediante una trasformazione lineare invertibile (capitolo XVI).[33] Esiste dunque una trasformazione

$$\begin{aligned} dy_1 &= a_{11}dx_1 + \cdots + a_{1n}dx_n\,, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ dy_n &= a_{n1}dx_1 + \cdots + a_{nn}dx_n\,, \end{aligned} \qquad (17)$$

tale che la sostituzione nella (13) dia

$$ds^2 = dx_1^2 + \cdots + dx_n^2\,.$$

Se cambiamo coordinate ponendo

$$\begin{aligned} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n\,, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_n &= a_{n1}x_1 + \cdots + a_{nn}x_n\,, \end{aligned}$$

i differenziali dy_j si esprimono mediante i dx_i proprio secondo le (17). Nel punto scelto, quindi, ds^2 assume la forma euclidea (16). La (13) definisce effettivamente una metrica riemanniana.

5. Il tensore di curvatura.

Lo spazio euclideo è il caso più semplice di spazio riemanniano.[34] Un compito fondamentale è descrivere analiticamente quanto una metrica riemanniana se ne discosti, definendo una misura della sua non euclideità. Tale misura è la curvatura.

Va subito chiarito che la curvatura non presuppone affatto che lo spazio si trovi in un altro spazio più grande, nel quale sarebbe piegato. È definita dall'interno ed esprime differenze rispetto alla geometria euclidea nelle proprietà intrinseche. Dire che lo spazio fisico ha curvatura significa attribuirgli proprietà geometriche non euclidee, non affermare che sia contenuto in uno spazio superiore. Una simile immagine non è richiesta dall'applicazione della geometria riemanniana e, se assunta come spiegazione necessaria, appartiene alla speculazione fantastica.

La curvatura riemanniana generalizza a n dimensioni la curvatura gaussiana delle superfici. Come abbiamo spiegato nel capitolo VII, §4, questa misura lo scostamento della geometria intrinseca dalla geometria piana e può essere definita interamente dall'interno. È precisamente la curvatura dello spazio riemanniano bidimensionale rappresentato dalla superficie.

Ricordiamo due formule. Su una superficie, vicino a O, consideriamo un piccolo triangolo con lati geodetici, angoli α, β, γ e area σ. La quantità $\alpha + \beta + \gamma - \pi$ misura lo scostamento della somma degli angoli dal valore euclideo.

Quando il triangolo si contrae verso *O*, il rapporto fra tale scostamento e l'area tende alla curvatura gaussiana *K* in *O*. Per un piccolo triangolo,

$$\frac{\alpha + \beta + \gamma - \pi}{\sigma} = K + \epsilon,$$

dove $\epsilon \to 0$ quando il triangolo si contrae verso *O*. La curvatura misura quindi precisamente lo scostamento della somma degli angoli dal valore piano.

Consideriamo ora una piccola circonferenza sulla superficie, centrata in *O*: il luogo dei punti a distanza intrinseca costante da *O*. Se *r* è il raggio e *l* la lunghezza, allora

$$l = 2\pi r - \frac{\pi}{3} K r^3 + \epsilon,$$

dove *K* è ancora la curvatura gaussiana in *O* e ϵ è un termine trascurabile rispetto a r^3 quando *r* tende a zero.

Qui la curvatura misura lo scostamento della lunghezza della circonferenza dal valore euclideo $2\pi r$.

In uno spazio riemanniano la curvatura ha un ruolo analogo. La si può definire fissando in *O* un piano bidimensionale di direzioni tangenti e formando una piccola superficie con le geodetiche uscenti da *O* in quelle direzioni. La curvatura gaussiana della superficie in *O* è detta curvatura sezionale dello spazio in quel piano. In generale cambia sia da punto a punto sia, nello stesso punto, al variare del piano. Non basta quindi un solo numero. Riemann trovò le relazioni fra queste curvature, mostrando che esse sono determinate da un sistema di quantità, le componenti del tensore di curvatura.

Non possiamo sviluppare qui l'apparato necessario. Importa comprendere che la curvatura è una misura intrinseca dello scostamento della metrica da quella euclidea. Governa, per esempio, le deviazioni della somma degli angoli da π e della lunghezza delle piccole circonferenze da $2\pi r$. In generale varia da punto a punto e, in uno stesso punto, richiede più componenti.

Uno spazio riemanniano può non essere omogeneo; allora non è possibile spostare liberamente le figure conservando tutte le distanze. Sorge quindi la domanda: in quali spazi è possibile una libera mobilità delle figure con gli stessi gradi di libertà dello spazio euclideo? Sono gli spazi con il massimo grado di omogeneità e isotropia.

Lo spazio euclideo è omogeneo e ha curvatura nulla. Un altro esempio è lo spazio di Lobačevskij. Entrambe le geometrie sono quindi casi particolari della geometria riemanniana.

La libera mobilità completa delle figure conduce a spazi di curvatura sezionale costante: il valore è lo stesso in ogni punto e in ogni piano di direzioni. La curvatura è allora caratterizzata da un unico numero. Localmente, curvatura zero significa geometria euclidea; curvatura costante negativa significa geometria di Lobačevskij; curvatura costante positiva significa geometria della sfera a *n* dimensioni nello spazio euclideo a $n + 1$ dimensioni. Per identificare

globalmente lo spazio con questi modelli occorrono ulteriori ipotesi, in particolare completezza e semplice connessione.

6. Applicazioni della geometria riemanniana.

Le applicazioni non tardarono. Riemann ne usò l'apparato formale in un problema di conduzione del calore; era soprattutto un'applicazione delle formule. L'idea di uno spazio astratto euclideo nell'infinitamente piccolo trovò invece un'applicazione nello spazio dei colori, dove la distanza fra colori vicini può essere modellata con una metrica riemanniana tridimensionale.

Un'altra applicazione importante nacque in meccanica. Consideriamo un punto materiale che si muova liberamente su una superficie liscia senza potersene staccare. Introduciamo coordinate x_1, x_2 sulla superficie: il moto è determinato dalla loro dipendenza dal tempo, e la velocità dalle derivate $\dot{x}_1$, $\dot{x}_2$. Il punto si muove dunque in uno spazio bidimensionale con una propria geometria intrinseca, in generale non euclidea. Le leggi del moto possono essere espresse mediante le due coordinate e le loro derivate prime e seconde.

Se agisce una forza, la componente normale viene bilanciata dalla reazione vincolare quanto necessario a mantenere il punto sulla superficie; il moto intrinseco è governato dalle componenti tangenziali. Possiamo quindi considerare le forze come agenti nella superficie stessa e il moto come un moto in uno spazio riemanniano bidimensionale. Le leggi hanno la stessa struttura di quelle ordinarie, ma tengono conto della metrica intrinseca. Lo chiarisce il fatto, già ricordato nel capitolo VII, §4, che un punto senza forze tangenziali e senza attrito percorre una geodetica a velocità scalare costante. Poiché le geodetiche svolgono il ruolo delle rette, è l'analogo della legge d'inerzia per un moto sulla superficie.

Fin qui il vantaggio dell'astrazione non è ancora evidente, perché stiamo descrivendo un moto su una superficie ordinaria.

Esso appare passando a sistemi meccanici la cui configurazione richiede più di due quantità. Non possiamo più rappresentarne il moto con un punto su una superficie. È la difficoltà già incontrata nel §7, che induce a sostituire i grafici ordinari con uno spazio multidimensionale.

Supponiamo che la configurazione di un sistema, cioè la disposizione delle sue parti, sia individuata da n quantità x_1, ..., x_n. Per un insieme di punti materiali liberi occorrono tre coordinate per punto. Un altro esempio è il giroscopio, una ruota che gira attorno a un asse a sua volta orientabile attorno a un punto fisso. La rotazione della ruota è descritta da un angolo, l'orientazione dell'asse da altri due: complessivamente tre quantità, almeno localmente, determinano la configurazione (figura 32).

Ogni configurazione può essere considerata come un punto dello spazio di tutte le configurazioni possibili, detto spazio delle configurazioni.[35] La sua dimensione è il numero delle quantità indipendenti necessarie, che ne costituiscono le coordinate. Per tre punti materiali liberi abbiamo nove coordinate; per il giroscopio tre coordinate angolari, quindi uno spazio tridimensionale.

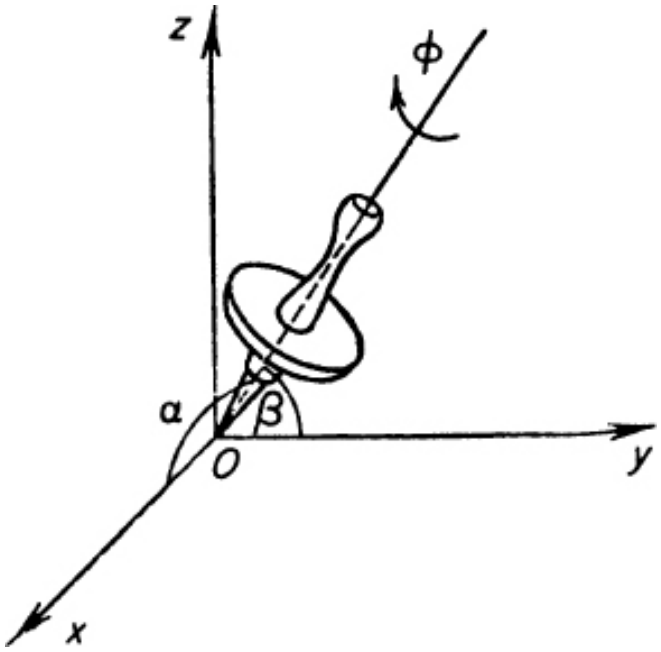


Figura 32.

Il moto del sistema è rappresentato dal moto di un punto nello spazio delle configurazioni. La velocità è determinata dalle velocità di variazione delle coordinate.

Torneremo su questi spazi nel capitolo XVIII per studiarne la struttura topologica. Qui osserviamo che vi si può introdurre una metrica strettamente legata alle proprietà meccaniche del sistema. Se l'energia cinetica è espressa da

$$T = \frac{1}{2} \sum_{i,k=1}^{n} a_{ik} \dot{x}_i \dot{x}_k ,$$

dove $\dot{x}_i$ sono le velocità di variazione delle coordinate, definiamo il quadrato dell'elemento di lunghezza mediante

$$ds^2 = \sum_{i,k=1}^{n} a_{ik}\, dx_i\, dx_k .$$

Lo spazio delle configurazioni diventa così riemanniano. Non soltanto il moto del sistema è rappresentato da quello di un punto: anche le equazioni del moto assumono la forma delle equazioni di un punto in quello spazio. La meccanica del sistema si traduce nella meccanica di un punto nello spazio delle configurazioni. In particolare, per sistemi con vincoli fissi e ideali del tipo considerato, il moto libero, come la rotazione libera di un giroscopio, corrisponde a un moto uniforme lungo una geodetica.

Questa rappresentazione, con varie estensioni e modifiche, è utile in numerosi problemi di meccanica teorica.

La geometria riemanniana è quindi un metodo di descrizione geometrica astratta dei fenomeni fisici. La descrizione non è arbitraria né un gioco intellettuale: riflette in forma astratta il meccanismo reale dei fenomeni, come ogni descrizione matematica della fisica. Lo

stesso vale per tutte le applicazioni della geometria astratta: cambiano la potenza e la finezza delle astrazioni, non la natura del procedimento.

L'applicazione più brillante si ebbe nella teoria della relatività. Ne parleremo nella sezione seguente, affrontando il difficile rapporto fra geometria astratta e proprietà dello spazio reale.

7. Generalizzazioni.

Nei trent'anni precedenti la stesura del libro, le geometrie non euclidee conobbero sviluppi e generalizzazioni notevoli. Nacquero teorie che comprendono la geometria riemanniana come caso particolare: anzitutto la geometria di Finsler, la cui idea risale allo stesso Riemann;[36] poi la teoria generale degli spazi di É. Cartan, che collega la geometria riemanniana al programma di Erlangen, e altre ancora. Non possiamo esporle qui; ricordiamo soltanto che si sviluppano soprattutto mediante apparati analitici specifici. Vi contribuirono anche geometri sovietici, fra i quali P. K. Raševskij con la geometria «polimetrica» e V. V. Vagner, le cui ricerche spaziavano dai problemi generali degli spazi curvi alle applicazioni meccaniche.

§10. Geometria astratta e spazio reale

1. La difficoltà di immaginare lo spazio reale come non euclideo.

Seguendo lo sviluppo delle idee geometriche da Lobačevskij in poi, ci siamo addentrati negli spazi astratti, allontanandoci dall'oggetto originario della geometria: lo spazio reale nel quale avvengono i fenomeni. Torniamo ora a esso per chiarire che cosa la geometria astratta abbia apportato alla conoscenza delle sue proprietà.

La geometria nacque dall'esperienza, dallo studio delle forme e delle relazioni spaziali dei corpi: misurazione dei terreni, capacità dei recipienti e così via. All'origine era quindi una teoria fisica, come la meccanica. Secondo la prospettiva dell'autore, gli assiomi euclidei esprimono in forma precisa risultati di una lunga esperienza e leggi della natura, così come si parla di assiomi della meccanica.[37] Ma non è lecito considerarli descrizioni fisiche assolutamente esatte, sottratte a ogni modifica o generalizzazione richiesta da nuovi dati: le proprietà dello spazio reale possono differire dalle previsioni euclidee.

Abbiamo già svolto questo ragionamento, che ora può sembrare evidente. Non lo era un secolo prima, quando le idee di Lobačevskij non avevano ottenuto riconoscimento generale. Prima di lui e di Gauss, l'eventuale inesattezza fisica della geometria euclidea non era stata considerata seriamente. Lobačevskij sviluppò la propria geometria come teoria delle possibili proprietà dello spazio reale; Riemann e altri ampliarono il problema, studiando leggi di misura che osservazioni più precise avrebbero potuto rivelare. In tal senso, alcune parti della geometria astratta sono teorie delle possibili strutture dello spazio. Queste restarono ipotesi finché la relatività generale di Einstein, nel 1915, diede forma fisica alle idee di Lobačevskij e Riemann, prevedendo una struttura spazio-temporale non euclidea e conseguenze osservabili su scala astronomica.

Queste considerazioni comportano almeno tre difficoltà. Chiarirle significa comprendere il rapporto fra geometria astratta e geometria fisica.

La prima è immaginare come e in quale senso lo spazio reale possa avere proprietà diverse da quelle euclidee. La consuetudine con queste ultime rende difficile concepire alternative.

La seconda riguarda l'espressione stessa «proprietà dello spazio reale». Tendiamo a immaginare lo spazio come vuoto e omogeneo, quasi che l'omogeneità fosse già contenuta nel concetto. Ma come può il «vuoto» avere proprietà? Usiamo abitualmente quell'espressione senza riflettere sul problema, che merita invece attenzione.

La terza riguarda la verità di una geometria. La risposta apparentemente semplice è: è vera quella che corrisponde alla realtà. Abbiamo però visto che i fatti geometrici interni a un cerchio possono essere descritti sia nel linguaggio euclideo sia nel modello di Lobačevskij. Entrambe le teorie corrispondono quindi, in un certo senso, alla realtà. Quale è vera, e in quale senso? Perché diciamo comunque che nel cerchio vale la geometria euclidea e che quella di Lobačevskij vi è soltanto rappresentata?

Sono questioni difficili, che hanno talvolta tratto in inganno anche matematici eminenti.

Cominceremo dalla seconda: capire che cosa siano le proprietà dello spazio aiuterà a sciogliere anche le altre difficoltà.

2. Spazio e materia.

Nell'interpretazione qui sostenuta dall'autore, le proprietà dello spazio sono proprietà dei corpi reali, delle loro forme e relazioni materiali. Un luogo, un punto o una direzione sono individuati mediante corpi. «Qui» e «là» hanno significato rispetto a qualcosa: sulla Terra, in questa stanza, e così via. Un luogo è sempre determinato mediante criteri materiali. Analogamente, una retta si manifesta come filo teso, bordo di un regolo o raggio luminoso. La retta astratta esprime proprietà comuni di queste realizzazioni, come il concetto di «casa» esprime proprietà comuni delle case concrete senza esistere come oggetto separato da esse.

L'autore collega questo carattere oggettivo delle proprietà spaziali alla tesi del materialismo dialettico secondo cui lo spazio è una forma di esistenza della materia. La forma di un oggetto dipende dai rapporti fra le sue parti; la struttura spaziale esprime regolarità comuni nelle relazioni fra corpi ed eventi: posizione, distanza, ordine. Come la forma non si separa dal contenuto se non nell'astrazione, così lo spazio non si separerebbe dalla materia. L'idea di uno spazio «in sé», privo di materia, è un'astrazione della quale non si deve abusare. «Fra», «dentro», «retto», «sferico» descrivono relazioni o forme dei corpi. La geometria le considera astrattamente per coglierne la generalità, ma non si deve sostituire il concetto astratto alla realtà oggettiva.

Proseguendo questa argomentazione filosofica, in uno spazio assolutamente privo di materia nulla distinguerebbe luoghi o direzioni: esso si ridurrebbe al nulla. Anche quando immaginiamo uno spazio vuoto, presupponiamo tacitamente che vi si possano distinguere luoghi e direzioni. Nell'astrazione conserviamo quindi proprietà che, secondo questa concezione, appartengono allo spazio reale per il suo legame con i corpi materiali.

In tale prospettiva lo spazio è una forma di esistenza della materia e le sue proprietà sono proprietà delle relazioni fra corpi: posizioni reciproche, dimensioni e così via.

Inoltre, per attribuire significato fisico ai teoremi geometrici occorre precisare che cosa rappresentino «retta», «distanza» e gli altri concetti. Nel §4 abbiamo visto che una stessa teoria ammette interpretazioni differenti.

Per confrontare la geometria con l'esperienza bisogna dunque definire il significato fisico dei concetti con la massima precisione possibile. Senza di esso, i teoremi hanno carattere matematico astratto e formale. Secondo l'autore, la seconda difficoltà nasce proprio dal sostituire allo spazio reale, legato alla materia, uno spazio «puro», che è soltanto un'astrazione.

3. Intuizione e comprensione.

Possiamo ora affrontare le altre due difficoltà.

Come immaginare proprietà non euclidee dello spazio reale? Per verificare, per esempio, che la somma degli angoli di un triangolo sia 180°, dobbiamo stabilire fisicamente che cosa siano i lati e come misurare gli angoli. Se i lati sono raggi luminosi nel vuoto, non è inconcepibile che misure molto precise diano una somma diversa. Analogamente, misurare raggio e circonferenza con la stessa unità potrebbe non confermare esattamente $l = 2\pi R$. Sulla Terra, per esempio, una circonferenza geodetica non ha lunghezza proporzionale al raggio geodetico: cresce fino al massimo quando il raggio raggiunge un quarto della lunghezza di un meridiano completo, poi decresce. Si può obiettare che sulla sfera i «raggi» sono archi di cerchio massimo e che ciò non contraddice la geometria euclidea dello spazio ambiente. Nella relatività generale, però, anche il rapporto fra circonferenza e raggio misurato mediante distanze proprie può differire da 2π. Per un corpo sferico omogeneo, nell'approssimazione di campo debole e con le convenzioni del modello, l'originale riporta la formula

$$\text{lunghezza della circonferenza / raggio} = 2\pi(1 - kM/(3Rc^2)),$$

dove M è la massa in tonnellate metriche, R il raggio in chilometri, c = 300.000 km/s la velocità della luce e k la costante gravitazionale, indicata nell'originale, con queste unità e l'approssimazione storica, come $66{,}6 \cdot 10^{-18}$.

Il rapporto risulta quindi un poco minore di 2π. I calcoli riportati nell'edizione originale danno uno scostamento di circa 0,000004 per la superficie del Sole e di circa 0,00014 per la compagna di Sirio, la cui densità media viene allora stimata in 50.000 volte quella dell'acqua. Questi numeri rispecchiano i dati e le approssimazioni del testo storico.

Si può ancora obiettare che tutto ciò non sia visualizzabile e che l'immagine intuitiva dello spazio resti euclidea. Ma il compito della scienza è comprendere i fenomeni, non necessariamente offrirne immagini familiari. L'intuizione visiva dipende dai sensi e dall'esperienza: non visualizziamo direttamente raggi ultravioletti, onde radio o il comportamento di un elettrone atomico, se non mediante modelli. Questo non impedisce di comprenderli, come mostrano, per esempio, i risultati della radiotecnica. Inoltre l'immaginazione si educa. Oggi immaginiamo senza difficoltà gli antipodi, dove le persone appaiono dal nostro punto di vista «a testa in giù»; un tempo la loro presunta inconcepibilità serviva come argomento contro la sfericità della Terra.

4. Geometria e verità.

Torniamo alla domanda su quale geometria sia vera. Gli stessi fatti interni a un cerchio possono essere espressi nei linguaggi di Euclide e di Lobačevskij; entrambe le descrizioni corrispondono a quei fatti. Non c'è da stupirsi: uno stesso fenomeno si descrive in modi diversi, una grandezza si misura in unità diverse, una curva ha equazioni diverse in sistemi di coordinate diversi. Così una raccolta particolare di relazioni geometriche può ammettere descrizioni differenti. Ma il problema fisico riguarda l'insieme generale delle relazioni spaziali, non un dominio isolato artificialmente.

In questa prospettiva i fatti geometrici non si separano dagli altri fenomeni cui sono legati. La lunghezza di un segmento, misurata riportandovi un regolo rigido, è connessa al movimento dei corpi e quindi alla meccanica. Nel modello circolare di Lobačevskij, invece, la misura avviene diversamente: una corda risulta infinita. Questa misura non coincide con quella nata dal confronto meccanico dei corpi reali. La congruenza euclidea deriva dalla possibilità di sovrapporre figure mediante movimenti rigidi; nel modello non euclideo il ruolo dei movimenti è assunto da altre trasformazioni. Quando si includono le relazioni meccaniche ordinarie, nel piccolo cerchio fisico vale dunque, con grande approssimazione, la geometria euclidea.

Nella geometria euclidea i movimenti corrispondono ai movimenti meccanici ordinari dei corpi rigidi. Per questo, sostiene l'autore, essa fu scoperta per prima. Lo sviluppo della fisica ha però mostrato che le leggi newtoniane e la descrizione euclidea sono approssimazioni di leggi più generali. Nel passaggio alla geometria della relatività hanno avuto un ruolo decisivo non soltanto la meccanica, ma anche l'elettromagnetismo e l'ottica. La geometria intesa come scienza dello spazio fisico dipende quindi dalla fisica e può esserne separata soltanto astrattamente e in contesti delimitati.

Lobačevskij indicò chiaramente questa dipendenza, prevedendo che nuove classi di fenomeni potessero richiedere leggi geometriche diverse. L'autore contrappone a tale posizione materialistica il convenzionalismo di Poincaré, secondo cui la scelta geometrica dipenderebbe dalla semplicità o dall'«economia del pensiero», per usare l'espressione di Mach. Poincaré riteneva che si sarebbe rinunciato più facilmente alla propagazione rettilinea della luce che alla geometria euclidea. Morì però tre anni prima della formulazione definitiva della relatività generale, nella quale si abbandona la geometria euclidea conservando, in forma generalizzata, una legge geometrica di propagazione della luce: essa segue geodetiche nulle dello spazio-tempo.

Possiamo così distinguere i due aspetti. Una collezione isolata di fatti può ammettere varie descrizioni valide. Ma per stabilire le proprietà fisiche dello spazio, la geometria va confrontata con l'intero insieme dei fenomeni. Collegandola alla fisica emergono differenze sostanziali fra teorie che, considerate soltanto formalmente, potrebbero essere confrontate per semplicità. Secondo l'autore, la geometria euclidea si impose perché corrispondeva alla meccanica ordinaria; lo sviluppo della relatività ha invece richiesto una geometria più complessa.

Rispetto allo spazio reale, una geometria è dunque fisicamente adeguata se descrive con sufficiente precisione l'insieme delle relazioni spaziali e concorda con la meccanica e con il resto della fisica, non soltanto con fatti isolati.

5. Lo spazio-tempo della relatività.

Quanto abbiamo detto non tocca ancora il contenuto centrale della teoria, che ha portato la comprensione di spazio e tempo oltre le prospettive di Lobačevskij e Riemann.

La sua idea fondamentale è che spazio e tempo non siano indipendenti, ma costituiscano una struttura unitaria: lo spazio-tempo quadridimensionale. Un evento è individuato da luogo e istante, quindi da tre coordinate spaziali e una temporale. Gli eventi formano così una varietà a quattro dimensioni. La relatività ne studia soprattutto la struttura, astraendo dalle particolarità dei singoli fenomeni. Non è soltanto una teoria dei moti rapidi o della cosmologia, né una teoria separata dello spazio o del tempo: è una teoria della loro struttura unitaria.

Anche nella meccanica newtoniana possiamo riunire formalmente spazio e tempo in quattro coordinate. Come abbiamo ricordato, Lagrange aggiungeva già t alle coordinate x, y, z di un punto materiale. Il moto diventa allora una linea in uno spazio a quattro coordinate. Questa unificazione, però, non stabilisce ancora il legame universale fra spazio e tempo introdotto dalla relatività. Le leggi di un singolo moto dipendono da posizione e tempo, ma nell'impostazione classica l'ordine spaziale e quello temporale restano nettamente separati. Successione e durata degli eventi vengono considerate assolute, indipendenti dal sistema di riferimento: domina il concetto di tempo assoluto.

La grande scoperta di Einstein fu che il tempo assoluto non descrive la realtà fisica. Poco dopo la teoria del 1905,[38] Minkowski mostrò che l'essenziale non era soltanto rinunciare a quel concetto, ma riconoscere il legame fra spazio e tempo in una struttura unitaria. La separazione fra coordinate spaziali e coordinata temporale dipende dal sistema di riferimento rispetto al quale si descrivono i fenomeni. Eventi simultanei in un sistema possono non esserlo in un altro.

L'ordine degli eventi non è però completamente arbitrario. Se due eventi possono essere collegati causalmente, il loro ordine causale è preservato dai cambiamenti di riferimento che rispettano l'orientazione temporale: la causa non segue l'effetto. Per eventi separati spazialmente, cioè non collegabili da un segnale luminoso o più lento, l'ordine temporale può invece dipendere dal riferimento. Poiché una configurazione spaziale riguarda eventi simultanei e la simultaneità è relativa, anche la separazione dello spazio dal complesso spazio-temporale dipende dal riferimento. Lo spazio è, in questo senso, una sezione dello spazio-tempo attraverso eventi simultanei nel sistema scelto.

Non possiamo esporre qui i fondamenti della relatività. Cercheremo soltanto di caratterizzarne alcuni aspetti nel linguaggio naturale della geometria astratta. Questa interpretazione, osserva l'autore, differisce sensibilmente dall'esposizione originaria di Einstein.

Possiamo considerare l'universo come un insieme di eventi. Un evento non è qui un fenomeno esteso nello spazio e nel tempo, ma un accadimento idealizzato come istantaneo e puntiforme, per esempio un lampo brevissimo di una sorgente puntiforme. Nel linguaggio geometrico, gli eventi sono punti della varietà quadridimensionale dell'universo.

Nella prospettiva dell'autore, lo spazio-tempo è la forma di questa varietà di eventi e la sua geometria è una struttura generale del mondo, determinata da relazioni universali fra eventi.

Anzitutto, come già per le relazioni spaziali, l'autore intende queste connessioni come materiali; le relazioni spaziali e temporali considerate in forma «pura» sono astrazioni.

Inoltre, le relazioni che determinano la struttura spazio-temporale devono avere carattere universale, come lo spazio-tempo stesso.

Una relazione fondamentale è quella causale: la possibilità che un evento influisca su un altro. L'autore propone di assumere questa possibilità di azione come base per comprendere la struttura spazio-temporale.

In questa interpretazione, lo spazio-tempo è l'insieme degli eventi considerati indipendentemente dalle loro particolarità, conservando la relazione generale di possibile influenza causale. Anche questa relazione viene astratta dalle forme concrete delle interazioni.

Nella relatività ristretta lo spazio-tempo è massimamente omogeneo. Esistono molte trasformazioni che preservano le relazioni causali. Consideriamo, per esempio, due coppie di eventi A, B e A', B', ciascuna separata spazialmente. Una trasformazione che preserva soltanto la struttura causale, ammettendo anche un cambiamento globale di scala, può portare la prima coppia nella seconda. Per ogni coppia X, Y viene conservata la possibilità o impossibilità di influenza causale fra le immagini X', Y'.

Dal punto di vista della relatività ristretta, lo spazio-tempo è quindi uno spazio quadridimensionale con una geometria caratterizzata da un gruppo di trasformazioni. Le trasformazioni che conservano l'intervallo spazio-temporale sono quelle di Lorentz e, includendo le traslazioni, quelle di Poincaré. Esse preservano anche la struttura causale e le leggi fisiche relativistiche. La sola conservazione della causalità, invece, ammette anche dilatazioni e non determina la scala metrica: non va confusa con l'invarianza dell'intervallo. Questa prospettiva si collega al programma di Erlangen di Klein discusso nel §6.

6. Gravitazione e curvatura.

La relatività generale abbandona l'ipotesi che lo spazio-tempo sia globalmente omogeneo: esso è approssimato localmente dallo spazio-tempo della relatività ristretta, ma può avere una struttura curva su regioni estese. Secondo Einstein, la geometria è legata alla distribuzione e al moto della materia e dell'energia; a sua volta determina il moto libero dei corpi. Questo legame si manifesta come gravitazione universale. La relatività generale è dunque una teoria della gravitazione che connette la struttura spazio-temporale alla dinamica della materia.

L'approssimazione omogenea in piccole regioni ricorda l'idea di spazio riemanniano euclideo nell'infinitamente piccolo. Lo spazio-tempo della relatività generale è infatti descritto da una geometria affine a quella riemanniana, ma con una differenza essenziale.

In uno spazio riemanniano quadridimensionale si possono scegliere coordinate tali che, nel punto considerato, il quadrato dell'elemento di lunghezza sia $ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2$.

Nello spazio-tempo si possono invece scegliere coordinate x, y, z, t per le quali, nell'evento considerato, l'elemento di intervallo sia $ds^2 = dx^2 + dy^2 + dz^2 - c^2dt^2$, dove c è la velocità della luce, riducibile a 1 scegliendo opportunamente le unità. Le prime tre coordinate sono spaziali e t è temporale. Il segno negativo del termine temporale esprime la differenza fondamentale fra tempo e spazio. La metrica è quindi lorentziana, o pseudo-riemanniana, non definita positiva.

La curvatura dello spazio-tempo svolge un ruolo centrale nella gravitazione. Le equazioni di Einstein collegano grandezze che descrivono la curvatura a quelle che descrivono la distribuzione e il moto della materia e dell'energia. Sono le equazioni del campo gravitazionale; da esse e dalle condizioni appropriate si ricavano anche le leggi del moto dei corpi. L'originale richiama in questo contesto i lavori di Einstein e di V. A. Fok, che contribuirono a sviluppare il problema.

Nella relatività generale la struttura spazio-temporale è complessa e la separazione fra spazio e tempo è meno semplice che nella teoria ristretta. Sotto determinate ipotesi e approssimazioni si può comunque considerare una geometria dello spazio. Essa è approssimativamente euclidea in regioni piccole rispetto alle scale di curvatura, mentre su regioni più grandi possono emergere differenze. Queste dipendono dalla distribuzione e dal moto delle masse e diventano misurabili vicino a corpi molto compatti o su scale cosmiche. Alcuni modelli dell'universo assumono una distribuzione mediamente uniforme della materia. In una delle soluzioni cosmologiche studiate da Fridman, le sezioni spaziali hanno geometria di Lobačevskij.

Nella relatività la geometria astratta non fornisce soltanto formule: il concetto stesso di spazio astratto permette di formulare più profondamente la teoria. Possibilità elaborate matematicamente trovano riscontro nei fenomeni, mostrando la forza del pensiero teorico. La geometria, nata dallo studio sperimentale delle forme e delle relazioni dei corpi, torna così allo spazio reale con strumenti sviluppati. È un percorso generale della scienza: dall'esperienza alle generalizzazioni e alle astrazioni, poi di nuovo all'esperienza, per comprendere più a fondo i fenomeni. Spiegando fatti noti e prevedendone di nuovi, la teoria orienta l'attività pratica e vi trova conferma e nuove ragioni di sviluppo.

Letture consigliate

H. S. M. Coxeter, Introduction to geometry, Wiley, New York, 1961.

L. P. Eisenhart, An introduction to differential geometry with use of the tensor calculus, Princeton University Press, Princeton, N. J., 1940.

--Riemannian geometry, Princeton University Press, Princeton, N. J., 1949.

W. C. Graustein, Elementary differential geometry, Macmillan, New York, 1935.

D. Hilbert, Foundations of geometry, Open Court, LaSalle, III., 1959.

A. V. Pogorelov, Differential geometry, Noordhoff, Groningen, 1959.

G. Y. Rainich, Mathematics of relativity, Wiley, New York, 1950.

D. J. Struik, Lectures on classical differential geometry, Addison-Wesley, Cambridge, Mass., 1950.

J. L. Synge, Relativity: the special theory, Interscience, New York, 1956.

Note al capitolo XVII

[1] Nota dell'edizione inglese: altra traslitterazione del nome, Nasiraddin.

[2] Furono formulate molte proposizioni equivalenti al quinto postulato. Per esempio: una parallela a una retta mantiene distanza costante da essa (Proclo); esistono triangoli simili ma non congruenti, con angoli uguali e lati diversi (Wallis); esiste almeno un rettangolo, cioè un quadrilatero con quattro angoli retti (Saccheri); una perpendicolare a un lato di un angolo acuto, innalzata su quel lato, incontra anche l'altro lato (Legendre); la somma degli angoli di un triangolo è due retti (Legendre); esistono triangoli di area arbitrariamente grande (Gauss). L'elenco potrebbe continuare.

[3] Lobačevskij scrisse queste parole nel corso di geometria del 1823, non pubblicato durante la sua vita. Il testo, intitolato *Geometria*, fu pubblicato per la prima volta nel 1910.

[4] Le «altre» proposizioni fondamentali della geometria vengono formulate precisamente nel §5.

[5] Se h è la lunghezza della perpendicolare e γ l'angolo di parallelismo, Lobačevskij dimostrò che $\tan(\gamma/2) = e^{-h/k}$, dove k è una costante di lunghezza ed e la base dei logaritmi naturali. Quando $h \to \infty$, il secondo membro tende a zero e quindi anche γ tende a zero.

[6] Se a, b sono i cateti e c l'ipotenusa di un triangolo rettangolo, al posto del teorema di Pitagora vale

$$2(e^{c/k} + e^{-c/k}) = (e^{a/k} + e^{-a/k})(e^{b/k} + e^{-b/k}).$$

Sviluppando in serie si ottiene

$$c^2 + \frac{c^4}{12k^2} + \cdots = a^2 + b^2 + \frac{a^4 + 6a^2b^2 + b^4}{12k^2} + \cdots,$$

Per k grande rispetto ai lati si ritrova approssimativamente il teorema di Pitagora, $c^2 = a^2 + b^2$. Inoltre, dalla formula $\tan(\gamma/2) = e^{-h/k}$, se h/k è piccolo segue $\tan(\gamma/2) \approx 1$ e quindi $\gamma \approx 90°$. A piccole distanze, le parallele del piano di Lobačevskij differiscono dunque poco da quelle euclidee.

[7] La storia fu in realtà ancora più complessa. Lobačevskij disponeva già degli strumenti per dimostrare la coerenza relativa mediante un modello analitico, ma non completò la dimostrazione, ottenuta più tardi. Inoltre Riemann propose nel 1854 una teoria, descritta nel §9, che conteneva implicitamente i risultati di Beltrami, senza formularli in modo esplicito. Il suo lavoro non fu compreso e venne pubblicato postumo nel 1868, lo stesso anno di quello di Beltrami. L'intera vicenda, dai tentativi di dimostrare il postulato alla comprensione della geometria non euclidea, mostra quanti sforzi e percorsi indiretti possano occorrere per scoprire una verità che poi appare semplice.

[8] La pseudosfera ha curvatura gaussiana costante negativa. Tutte le superfici con la stessa curvatura costante negativa hanno localmente la stessa geometria intrinseca e possono quindi rappresentare parti del piano di Lobačevskij. Hilbert dimostrò nel 1901 che non si può realizzare l'intero piano iperbolico completo come superficie regolare di questo tipo nello spazio euclideo tridimensionale. I risultati di Kuiper del 1955 consentono invece realizzazioni isometriche di regolarità C^1: esiste il piano tangente continuo, ma non la regolarità seconda necessaria a definire la curvatura classica della superficie. La diversa regolarità evita quindi il contrasto con il teorema di Hilbert.

Se la geometria di Lobačevskij è rappresentata su una superficie di curvatura costante negativa K, la costante k delle formule precedenti soddisfa $k^2 = -1/K$.

[9] Siano *A*, *B* sulla corda di estremi *C*, *D*, nell'ordine *C*, *A*, *B*, *D* (figura 12). Si forma, con lunghezze euclidee, il birapporto $(CB/CA)/(DB/DA)$. La distanza iperbolica è proporzionale al suo logaritmo; con la costante k usata nel testo vale $(k/2)\ln[(CB \cdot DA)/(CA \cdot DB)]$. Cambiando unità si può assumere il logaritmo stesso come misura.

[10] Si dice che la geometria di Lobačevskij è rappresentabile entro quella euclidea: la sua coerenza segue quindi dalla coerenza della geometria euclidea.

[11] Nel modello di Poincaré il piano è ancora l'interno di un cerchio, ma le rette sono archi di circonferenze ortogonali al bordo, oppure diametri. I movimenti sono automorfismi conformi del disco e, includendo quelli che invertono l'orientazione, trasformazioni anticonformi. Il legame con le trasformazioni conformi conduce alla teoria delle funzioni di variabile complessa.

[12] Si può sostituire «destra» con «sinistra».

[13] Se D_1 è espresso dalle (3) del §4, il movimento inverso è espresso dalle (4).

[14] Una trasformazione biunivoca del cerchio aperto in sé che porta rette in rette si estende univocamente, conservando questa proprietà, all'intero piano proiettivo, cioè al piano completato con i punti all'infinito.

[15] Se le tangenti in *B* e *C* sono parallele, il loro punto d'incontro *D* è un punto all'infinito del piano proiettivo.

[16] Poncelet, ingegnere militare francese, svolse le ricerche geometriche durante la prigionia in Russia dopo il 1812. Il suo *Traité des propriétés projectives des figures* apparve nel 1822.

[17] «Gruppo» non significa soltanto collezione. Un gruppo di trasformazioni (capitolo XX) contiene l'identità, che lascia fissi tutti i punti; insieme a ogni trasformazione contiene l'inversa, che riporta i punti nelle posizioni iniziali; insieme a due trasformazioni contiene la loro composizione, equivalente a eseguirle successivamente.

[18] Se $z = \rho(\cos\phi + i\sin\phi)$, sappiamo dal capitolo IV, §3, che

$$z^n = \rho^n(\cos\phi n + i\sin\phi n);$$

Quindi, quando l'argomento ϕ di z varia da 0 a 2π, l'argomento di z^n varia da 0 a $2\pi n$: il raggio vettore di z^n compie n giri completi.

[19] La generalizzazione comprende anche le riflessioni rispetto a iperpiani, poiché anch'esse conservano le distanze, oltre al passaggio a n variabili.

[20] Non si possono scegliere arbitrariamente quali siano le k variabili libere. Nel sistema $x_1 + x_2 + x_3 = 0$, $x_1 - x_2 - x_3 = 0$, per esempio, x_1 è necessariamente zero e non può fungere da parametro libero per esprimere le altre. Si afferma soltanto che una scelta adatta di k variabili esiste sempre.

[21] Nota dell'edizione inglese: una sfera può quindi essere rappresentata come punto di uno spazio a quattro dimensioni.

[22] Ogni scelta di m vertici di un simplesso determina una sua faccia, il simplesso di dimensione $m - 1$ generato da quei vertici. Il numero di tali facce è perciò il numero di combinazioni di m elementi fra $n + 1$:

$$C_m^{n+1} = \frac{(n+1)!}{m!\,(n-m+1)!}.$$

[23] Si cercano i poliedri convessi che tassellano lo spazio mediante traslazioni, con incontri lungo facce intere. Fedorov pose e risolse il problema in tre dimensioni per le esigenze della cristallografia. Voronoj e i successori progredirono in dimensione arbitraria; l'originale riferisce che una soluzione completa era allora nota soltanto nelle dimensioni due, tre e quattro. La formulazione riflette lo stato delle conoscenze dell'epoca.

[24] Il volume della sfera piena si calcola anche con la (9): una sezione di una palla a n dimensioni è una palla a $n - 1$ dimensioni. Si può quindi procedere ricorsivamente dalla dimensione $n - 1$ alla dimensione n.

[25] Si parla qui della percezione cromatica, non dello spettro della luce. La percezione è una risposta alla luce e stimoli spettralmente diversi possono produrre la stessa sensazione cromatica. Un esempio è il giallo ottenuto con luce gialla oppure con un'opportuna miscela additiva di rosso e verde. Nell'originale compare erroneamente il verde ottenuto mescolando rosso e blu. Le alterazioni della visione cromatica non sono tutte uguali: nelle dicromazie restano due componenti funzionali indipendenti; nelle rare monocromazie ne resta una. Non ogni forma di daltonismo è dicromatica.

[26] L'originale propone la miscela di polveri colorate fini a illuminazione costante. Per le relazioni lineari delle coordinate impiegate qui, tuttavia, l'interpretazione corretta è la miscelazione additiva di luci: le miscele di pigmenti non seguono in generale le stesse leggi lineari.

[27] Se C_0 e C_1 hanno coordinate (x_0, y_0, z_0) e (x_1, y_1, z_1), una miscela intermedia ha coordinate $x = (1 - t)x_0 + tx_1$, $y = (1 - t)y_0 + ty_1$, $z = (1 - t)z_0 + tz_1$, con $0 \le t \le 1$. I pesi sono t per C_1 e $1 - t$ per C_0.

[28] N. I. Lobačevskij, *Opere complete*, vol. II, Gostehizdat, 1949, p. 168.

[29] Si suppone che $f(x, y)$ sia continua nelle due variabili in un intorno del punto iniziale.

[30] Il significato preciso della (12) si esprime considerando una curva uscente da A, con coordinate $x_1(t)$, …, $x_n(t)$. Nel punto A, il differenziale ds della lunghezza d'arco è dato dalla (12).

[31] Se un unico sistema di coordinate rendesse euclidea la metrica in tutta la regione considerata, la geometria locale di quella regione sarebbe euclidea. Per identificarla globalmente con tutto lo spazio euclideo occorrono anche le appropriate condizioni globali.

[32] Una forma quadratica è un polinomio omogeneo di grado due nelle quantità considerate.

[33] Il fatto che le variabili siano differenziali non cambia il ragionamento algebrico: nel punto fissato le consideriamo come variabili indipendenti.

[34] In coordinate cartesiane ortogonali euclidee, la (16) dà $ds^2 = \sum dx_i^2$. Passando ad altre coordinate, il ragionamento del punto 4 esprime ds^2 mediante una forma quadratica come la (13). La formula generale vale quindi anche nello spazio euclideo. La particolarità di questo spazio è che esiste un unico sistema cartesiano nel quale la forma semplice vale ovunque, mentre in uno spazio riemanniano generale si può ottenerla soltanto punto per punto con scelte diverse di coordinate.

[35] Non va confuso con lo spazio delle fasi del §8, punto 2: in quest'ultimo un punto determina sia la configurazione sia le velocità istantanee del sistema.

[36] Paul Finsler avviò negli anni 1916-1918 lo studio sistematico della geometria qui menzionata.

[37] L'originale osserva che nei cinquant'anni precedenti si era sviluppata una concezione astratta degli assiomi, indipendente dal contenuto iniziale. L'autore sostiene che ciò non elimini il fondamento sperimentale della geometria fisica. Parlare di assiomi mette in primo piano la costruzione deduttiva: una proposizione è assunta come base della teoria e le altre, i teoremi, vengono ricavate mediante ragionamenti logici. Questa funzione formale va distinta dall'eventuale interpretazione della proposizione come legge empirica.

[38] La teoria del 1905 è detta relatività ristretta, o speciale, per distinguerla dalla relatività generale del 1915.

Nota editoriale

Sono conservate le dieci sezioni, le 32 figure numerate, le formule, la bibliografia e le 38 note. Il diagramma degli stati dell'acqua (figura 30) è ridisegnato con etichette italiane. Alcune formule sono ricomposte; sono corretti il segno del termine quadratico nello sviluppo dell'esponenziale negativo, gli indici delle costanti nei sistemi del §6 e quelli delle derivate parziali di f_1 nel §9. Il richiamo alla nota 4, spostato nella versione elettronica, è ricollocato nel testo.

Per evitare equivoci sono esplicitate alcune ipotesi matematiche: regolarità nelle rappresentazioni locali, convessità dei politopi regolari, misurabilità nelle definizioni di volume e condizioni di applicabilità dei teoremi di esistenza. Si completa l'elenco degli assiomi della chiusura topologica. La coerenza della geometria di Lobačevskij è relativa a quella dei modelli utilizzati; l'assenza di contraddizioni note, da sola, non ne costituisce una dimostrazione. L'uso storico della parola «iperpiano» è distinto da quello moderno.

Si distinguono proprietà locali e globali delle varietà: una geodetica minimizza la lunghezza solo localmente, l'immersione nella dimensione indicata richiede ipotesi di regolarità e va intesa localmente, mentre la classificazione globale degli spazi a curvatura costante richiede completezza e semplice connessione. Si precisa la differenza fra geometria riemanniana e pseudoriemanniana, fra stati quantistici e loro rappresentanti vettoriali e fra trasformazioni di Lorentz e trasformazioni più generali che conservano le relazioni causali.

Le considerazioni filosofiche restano attribuite all'autore; stime astronomiche, esempi fisici e riferimenti storici rispecchiano l'epoca del libro. Sono chiariti i limiti dell'esempio dei tre colori fondamentali e corretta la confusione fra miscelazione additiva delle luci e mescolanza di pigmenti. Nelle note sono distinti i risultati di Hilbert per superfici regolari da quelli di Kuiper in classe C^1. Queste precisazioni editoriali servono alla lettura del testo storico e non costituiscono un aggiornamento sistematico dell'opera.

CAPITOLO XVIII

Topologia

P. S. Aleksandrov

§1. L'oggetto della topologia

«Il contatto è la proprietà distintiva dei corpi e ci permette di chiamarli geometrici quando conserviamo in essi questa proprietà, facendo astrazione da tutte le altre, essenziali o accidentali che siano». Con queste parole N. I. Lobačevskij inizia il primo capitolo dei suoi *Nuovi principi della geometria*.[1]

Illustrando con un disegno (figura 1) le parole appena citate, Lobačevskij prosegue: «Due corpi *A*, *B* che si toccano formano un unico corpo geometrico *C*... Viceversa, ogni corpo *C* può essere diviso da una sezione arbitraria *S* in due parti *A*, *B*».

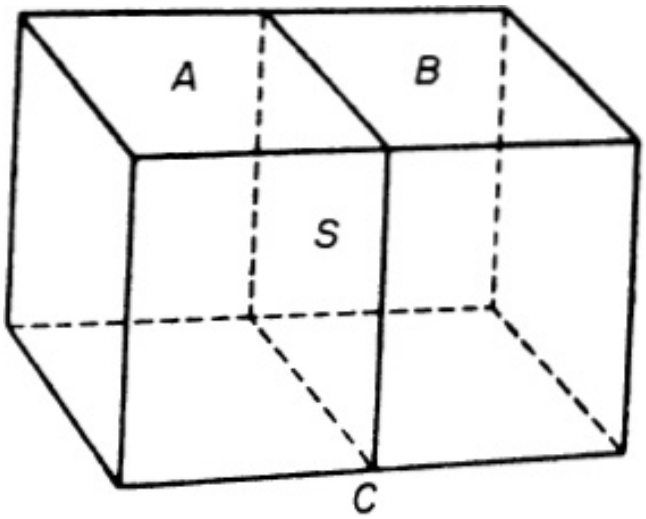


Figura 1.

I concetti di contatto, vicinanza, prossimità infinita, insieme al concetto di separazione di un corpo, che in un certo senso è loro duale, sono quelli che Lobačevskij pone a fondamento dell'intera costruzione della geometria. Essi sono anche, nella sostanza, i concetti fondamentali e originari della topologia in tutta l'ampiezza con cui oggi intendiamo questa disciplina. Hanno dunque ragione i commentatori contemporanei del grande geometra quando affermano[2] che «Lobačevskij compie il primo tentativo, nella storia delle scienze matematiche, di partire dalle proprietà topologiche dei corpi per costruire la geometria... I concetti di superficie, linea e punto sono definiti da Lobačevskij mediante le separazioni e i contatti fra corpi». I disegni della figura 2, tratti dalla sua opera, danno un'idea della varietà del contenuto geometrico concreto che egli vedeva riflesso nei concetti di contatto e di separazione dei corpi.

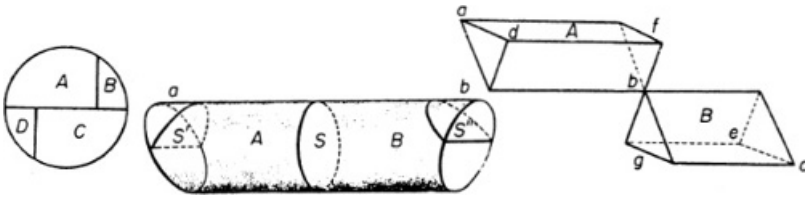


Figura 2.

Ogni trasformazione di una figura geometrica che non distrugga le relazioni di contatto fra le sue diverse parti si dice continua; se, inoltre, non crea nuovi contatti, la trasformazione si dice topologica. In una trasformazione topologica di una figura arbitraria, dunque, le parti che si toccano continuano a toccarsi e quelle che non si toccano non possono entrare in contatto. In breve, non possono prodursi né strappi né saldature. In particolare, due punti distinti non possono essere riuniti in un solo punto, perché in tal caso nascerebbe un nuovo contatto (figura 3). Considerando la figura come l'insieme dei punti che la costituiscono, una trasformazione topologica è perciò non soltanto continua, ma anche biunivoca: due punti distinti vengono trasformati in due punti distinti. Le trasformazioni topologiche sono quindi corrispondenze univoche e continue in entrambi i sensi.

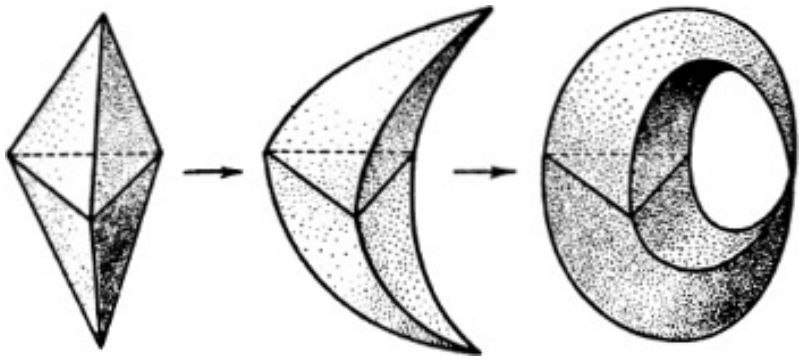


Figura 3.

Intuitivamente, una trasformazione topologica di una figura geometrica arbitraria - una curva, una superficie e così via - può essere rappresentata nel modo seguente.

Immaginiamo che la figura sia fatta di un materiale flessibile ed estensibile, per esempio di gomma. Possiamo sottoporla a ogni sorta di deformazioni continue, allungandone alcune parti e contraendone altre, e modificarne in ogni modo dimensioni e forma. Se, per esempio, un anello chiuso di gomma ha inizialmente la forma di una circonferenza, possiamo allungarlo fino a ottenere un'ellisse molto schiacciata, dargli la forma di un poligono regolare o irregolare e persino quella di curve chiuse assai complicate, alcune delle quali sono rappresentate nella figura 4. Non possiamo però trasformare topologicamente una circonferenza in una curva a otto: occorrerebbe saldare due suoi punti distinti (figura 5). Né possiamo trasformarla in un intervallo: dovremmo saldare una semicirconferenza all'altra oppure spezzare la circonferenza in un punto. La circonferenza è la più semplice delle curve chiuse: forma un solo cappio, mentre la curva a otto ne forma due e la curva a trifoglio della figura 5 ne forma tre. La proprietà della circonferenza di essere una curva chiusa semplice si conserva sotto ogni trasformazione topologica; è dunque, come si dice, una proprietà topologica.

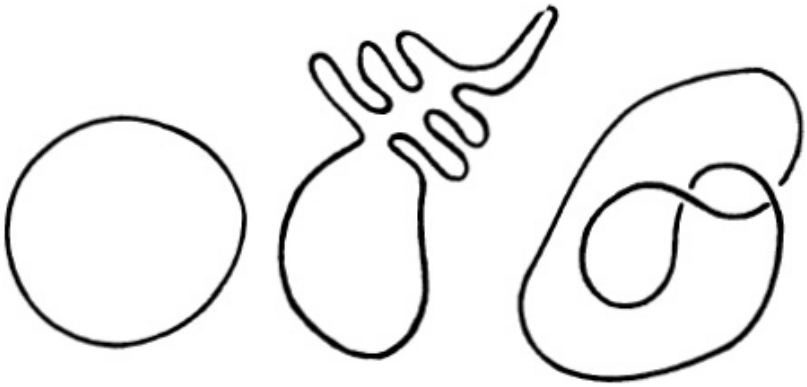


Figura 4.



Figura 5.

Prendiamo ora una superficie sferica, immaginandola come una sottile membrana di gomma. Anche in questo caso una trasformazione topologica può modificarne profondamente la forma (figura 6). Non può però trasformarla in un quadrato o in una superficie ad anello, come quella di un volante o di un salvagente, che si chiama toro (figura 7). La superficie sferica possiede infatti due proprietà che si conservano sotto ogni trasformazione topologica. In primo luogo è chiusa: non ha bordi, mentre il quadrato ne ha. In secondo luogo, ogni curva chiusa semplice tracciata sulla sfera la separa, nel senso di Lobačevskij: tagliando la membrana lungo tale curva, la superficie si divide in due parti disconnesse. Il toro non ha questa proprietà. Se lo tagliamo lungo un meridiano (figura 7), non si divide in due parti, ma diventa una superficie a forma di tubo ricurvo (figura 8), che possiamo facilmente raddrizzare e trasformare topologicamente in un cilindro. A differenza della sfera, dunque, non ogni curva chiusa semplice sul toro è una curva di separazione. Perciò la superficie sferica non può essere trasformata topologicamente in un toro. Diciamo che sfera e toro sono superfici topologicamente distinte, che appartengono a tipi topologici differenti o, infine, che non sono omeomorfe. Viceversa, una sfera, un ellissoide e, più in generale, la superficie chiusa di qualsiasi corpo convesso limitato appartengono allo stesso tipo topologico: sono omeomorfe. Ciò significa che si possono trasformare l'una nell'altra mediante una trasformazione topologica.

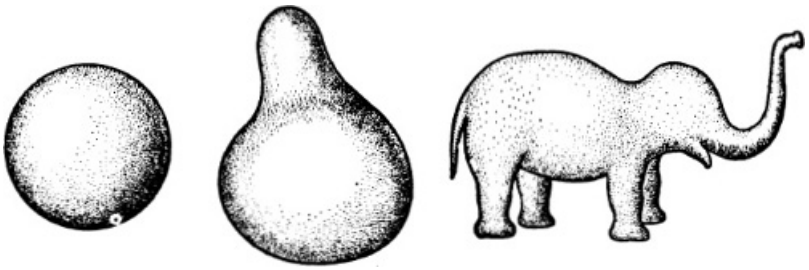


Figura 6.

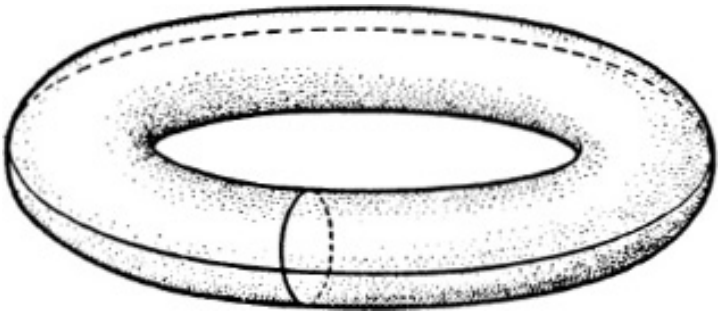


Figura 7.

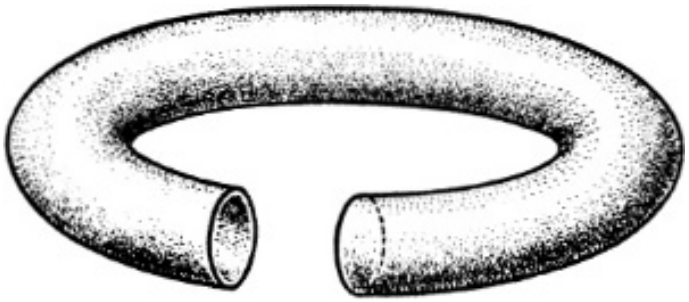


Figura 8.

§2. Le superfici

Come abbiamo detto, si chiama topologica ogni proprietà di una figura geometrica che si conserva sotto una sua trasformazione topologica arbitraria. La topologia studia le proprietà topologiche delle figure; studia inoltre le trasformazioni topologiche e, più in generale, le trasformazioni continue delle figure geometriche.

Abbiamo appena incontrato alcuni esempi di proprietà topologiche: per una curva o una superficie, l'essere chiusa; per una curva chiusa, l'essere semplice, cioè formata da un solo cappio; per una superficie, il fatto che ogni curva chiusa semplice tracciata su di essa la separi, come accade sulla sfera ma non sul toro; e così via.

Il massimo numero di curve chiuse che si possono tracciare su una superficie in modo che, tagliando lungo tutte queste curve, la superficie non si divida in parti, si chiama ordine di connessione della superficie. Questo numero fornisce informazioni importanti sulla sua struttura topologica. Abbiamo visto che per la sfera è zero: ogni curva chiusa semplice sulla sfera la separa. Sul toro possiamo invece trovare due curve chiuse che, prese insieme, non lo separano: un meridiano arbitrario e un parallelo (figura 7). Non possiamo però tracciarvi tre curve chiuse che, prese insieme, non producano una separazione; l'ordine di connessione del toro è 2. Quello della superficie a due manici della figura 9 è 4, e così via. In generale, prendiamo una superficie sferica e pratichiamovi $2p$ fori circolari: nella figura 10 è illustrato il caso $p = 3$. Dividiamo i fori in p coppie e colleghiamo i bordi di ogni coppia con un tubo cilindrico, cioè con un «manico». Otteniamo una sfera con p manici, detta anche superficie normale di genere p. Il suo ordine di connessione è $2p$.



Figura 9.

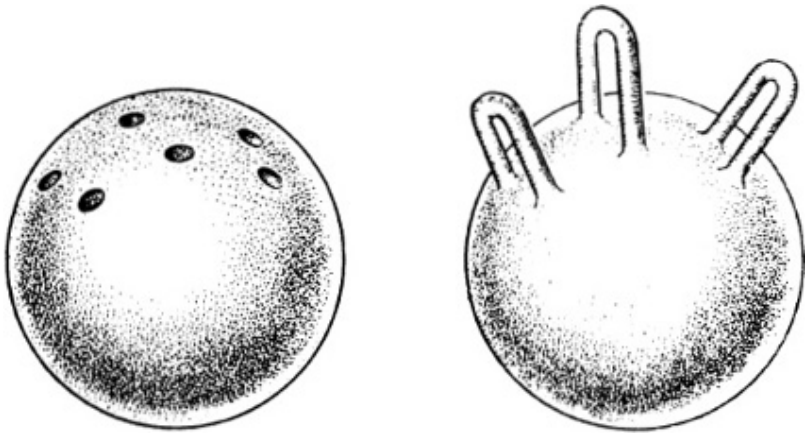


Figura 10.

Tutte queste superfici, nell'espressione di Lobačevskij, «separano» lo spazio: ciascuna lo divide in due regioni, una interna e una esterna, delle quali costituisce la frontiera comune. Questo fatto è collegato a un altro: ognuna delle nostre superfici ha due facce, una interna e una esterna, che possono essere dipinte con colori differenti.

Esistono però anche superfici dette unilaterali, che non possiedono due facce distinte. La più semplice è il noto nastro di Möbius. Per ottenerlo prendiamo una striscia rettangolare di carta *ABCD* e incolliamo i due lati corti opposti *AB* e *CD* facendo coincidere *A* con *C* e *B* con *D*. La superficie ottenuta, rappresentata nella figura 11, si chiama nastro o striscia di Möbius. È facile verificare che non ha due facce da colorare diversamente: percorrendo la linea mediana a partire, per esempio, dal punto *E*, torniamo in *E* sull'altra faccia della striscia, pur senza averne mai attraversato il bordo. Si osservi inoltre che il bordo del nastro di Möbius è costituito da un'unica linea chiusa.

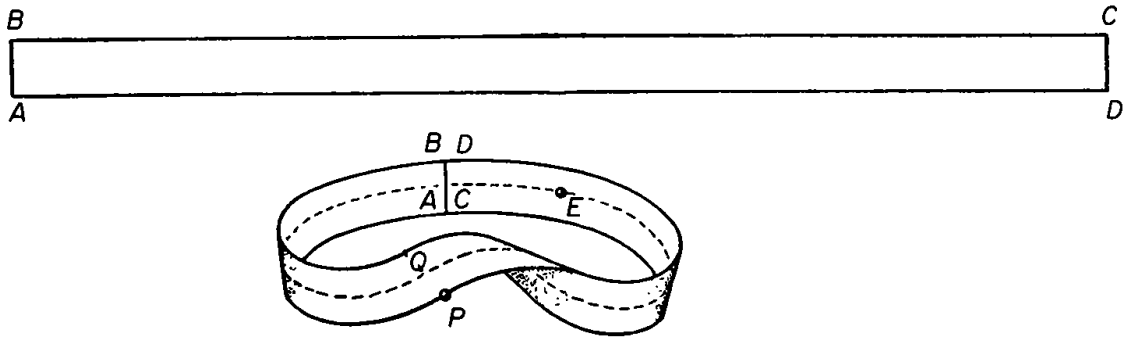


Figura 11.

Sorge allora una domanda: esistono superfici chiuse unilaterali, cioè superfici unilaterali senza bordo? Sì, ma qualunque loro rappresentazione nello spazio tridimensionale presenta necessariamente autointersezioni. Un esempio tipico è illustrato nella figura 12: il cosiddetto «toro unilaterale», o bottiglia di Klein. Se, senza preoccuparci delle autointersezioni, immaginiamo di incollare due nastri di Möbius lungo i loro bordi - ciascuno dei quali, come abbiamo detto, è un unico contorno - otteniamo una bottiglia di Klein.

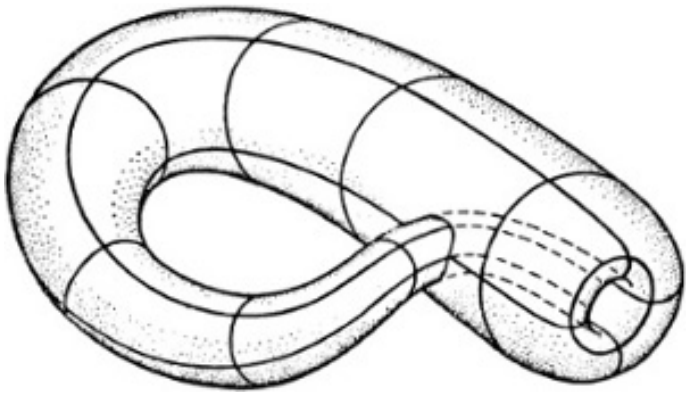


Figura 12.

Possiamo ora formulare il teorema fondamentale della topologia delle superfici nel caso bilaterale: ogni superficie chiusa, connessa e bilaterale è omeomorfa a una superficie normale di un certo genere p, ossia a una sfera con p manici. Due superfici di questo tipo sono omeomorfe se e solo se hanno lo stesso genere p, e quindi lo stesso ordine di connessione $2p$: devono cioè essere omeomorfe a una sfera con lo stesso numero di manici.

Anche per le superfici unilaterali esistono forme normali, analoghe alle superfici bilaterali di genere p, ma più difficili da descrivere. Occorre prendere una sfera, praticarvi p fori circolari e attaccare a ciascuno di essi un nastro di Möbius, incollandone il bordo al contorno del foro. La difficoltà di immaginare questo incollamento nasce dal fatto che non è possibile realizzarlo materialmente nello spazio ordinario senza autointersezioni: esse compaiono subito e sono inevitabili in ogni modello spaziale di una superficie chiusa unilaterale.

Non bisogna pensare che le superfici chiuse unilaterali siano semplici curiosità matematiche, prive di rapporti con i problemi seri della scienza. Per capire quanto sarebbe errata questa opinione basta ricordare che una delle conquiste fondamentali del pensiero geometrico è stata la creazione della geometria proiettiva, i cui elementi sono ormai insegnati nei corsi universitari e negli istituti di formazione degli insegnanti. Le sue origini pratiche risiedono nella teoria della prospettiva, sviluppatasi nel Rinascimento, con Leonardo da Vinci, per le esigenze dell'architettura, della pittura e del disegno tecnico. Nel Cinquecento e nel Seicento furono scoperti i primi teoremi della geometria proiettiva. Nata dunque da esigenze pratiche ben determinate, essa divenne, nel suo pieno sviluppo teorico, una delle più significative generalizzazioni della geometria. Fu proprio nell'ambito della geometria proiettiva che la geometria non euclidea di Lobačevskij ricevette per la prima volta una comprensione completa.[3]

Il passaggio dal piano ordinario della geometria elementare al piano proiettivo consiste nel completare il piano con nuovi elementi astratti, i punti impropri o «all'infinito». Solo dopo tale completamento la proiezione di un piano su un altro - per esempio quella prodotta da una lanterna da proiezione su uno schermo - diventa, nelle usuali condizioni non degeneri, una corrispondenza biunivoca fra i due piani. Questo completamento, che in geometria analitica corrisponde al passaggio dalle coordinate cartesiane alle coordinate omogenee, si effettua così. A ogni retta si aggiunge un unico punto improprio, e due rette hanno lo stesso punto improprio se e solo se sono parallele. Una retta completata con un punto all'infinito diventa

una linea chiusa. L'insieme dei punti impropri di tutte le rette forma, per definizione, la retta impropria, o retta all'infinito.

Poiché rette parallele hanno lo stesso punto all'infinito, per rappresentare l'intero processo di completamento basta considerare le rette passanti per un punto arbitrario del piano, per esempio per l'origine delle coordinate *O* (figura 13). I loro punti impropri esauriscono quelli dell'intero piano proiettivo: ogni retta ha infatti lo stesso punto improprio della parallela condotta per *O*. Otteniamo così un «modello» del piano proiettivo immaginando un disco di raggio «infinitamente grande», con centro in *O*, e identificando ogni coppia di punti diametralmente opposti *A*, *A′* della sua circonferenza in un unico punto «all'infinito» della retta *AA′*. La circonferenza diventa allora la retta all'infinito, purché si tenga sempre presente che ogni coppia di suoi punti diametralmente opposti deve essere considerata come un solo punto. Si comprende così che il piano proiettivo è una superficie chiusa, priva di bordo.

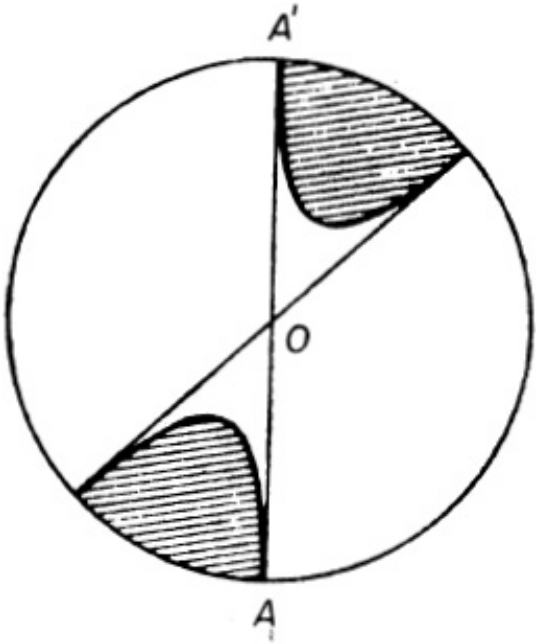


Figura 13.

Prendiamo nel piano proiettivo una curva del secondo ordine che nel piano ordinario si presenta come un'iperbole (figura 13). Nel piano proiettivo essa è una curva chiusa, divisa dalla retta all'infinito in due rami. Ricordando l'identificazione dei punti diametralmente opposti della circonferenza del disco fondamentale, vediamo facilmente che la regione interna all'iperbole, tratteggiata nella figura 13, è omeomorfa all'interno di un disco ordinario, mentre la regione complementare, non tratteggiata, è un nastro di Möbius se vi includiamo il bordo. Dal punto di vista topologico, dunque, il piano proiettivo si ottiene incollando lungo il bordo un disco - nel nostro caso la regione racchiusa dalla conica - e un nastro di Möbius. Ne segue che il piano proiettivo, oggetto fondamentale della geometria proiettiva piana, è una superficie chiusa unilaterale.

Oltre al suo grande valore geometrico intrinseco, l'esempio del piano proiettivo mette in luce una caratteristica del pensiero geometrico contemporaneo, sviluppatosi sulla base della scoperta di Lobačevskij. Il pensiero geometrico è sempre stato astratto, per la natura stessa

del concetto di figura geometrica. Qui raggiunge però un nuovo grado di astrazione, che si manifesta nel completamento del piano ordinario con nuovi elementi, i punti impropri. Anche questi elementi astratti hanno naturalmente una base reale: ogni «punto improprio» è l'astrazione di un fascio di rette parallele. Tuttavia li introduciamo come elementi geometrici distinti, che possiamo raffigurarci soltanto indirettamente, mediante la «fusione», non realizzabile fisicamente, di punti diametralmente opposti della circonferenza di un disco. Costruzioni astratte di questo genere hanno grandissimo valore in tutta la topologia contemporanea, soprattutto nel passaggio dalle superfici alle varietà di tre o più dimensioni.

§3. Le varietà

Consideriamo il semplice dispositivo talvolta chiamato pendolo doppio piano (figura 14). Esso è formato da due aste *OA* e *AB*, incernierate in *A*. Il punto *O* è fisso; l'asta *OA* ruota liberamente attorno a *O* in un piano fissato e l'asta *AB* ruota nello stesso piano attorno ad *A*. Ogni configurazione del sistema è completamente determinata dagli angoli ϕ e ψ che le due aste formano con una direzione fissa del piano, per esempio il verso positivo dell'asse delle ascisse. Possiamo considerare questi due angoli, variabili da 0 a 2π, come le «coordinate geografiche» di un punto del toro, misurate rispettivamente a partire dall'«equatore» e da uno dei «meridiani» del toro (figura 15).

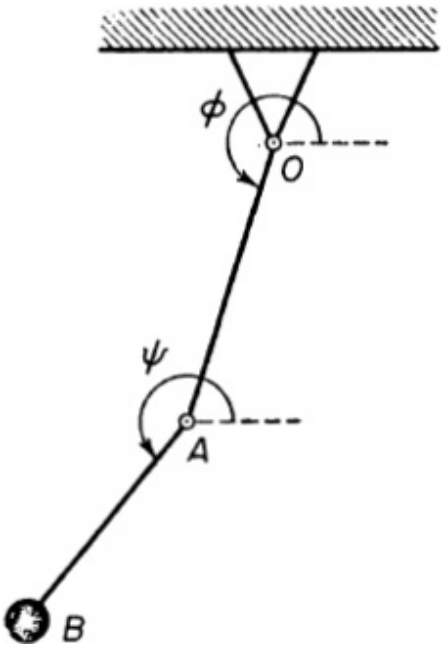


Figura 14.

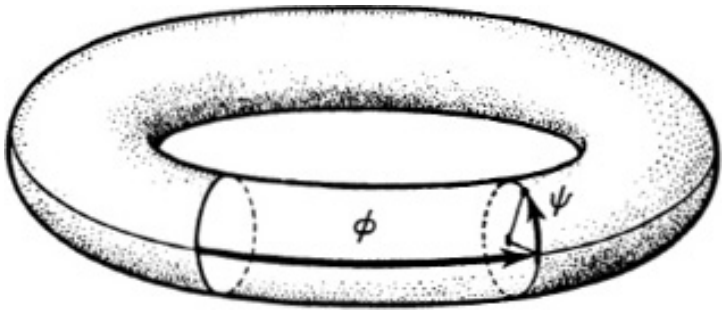


Figura 15.

Possiamo dunque dire che la varietà di tutte le configurazioni del nostro sistema meccanico ha due dimensioni ed è precisamente un toro. Se sostituiamo ciascuno dei due angoli ϕ, ψ con il punto corrispondente di una circonferenza sulla quale siano fissati un punto iniziale e un verso (figura 16), possiamo dire anche che ogni configurazione è completamente descritta da un punto su ciascuna di due circonferenze: uno rappresenta la latitudine ϕ, l'altro la longitudine ψ. Come in geometria analitica identifichiamo un punto del piano con una coppia di numeri, le sue coordinate, così qui identifichiamo un punto del toro - e quindi una configurazione del pendolo - con la coppia delle sue coordinate geografiche, ossia con una coppia di punti appartenenti a due circonferenze. Il contenuto essenziale di questa costruzione si esprime dicendo che la varietà di tutte le configurazioni del pendolo doppio piano, cioè il toro, è il prodotto topologico di due circonferenze.

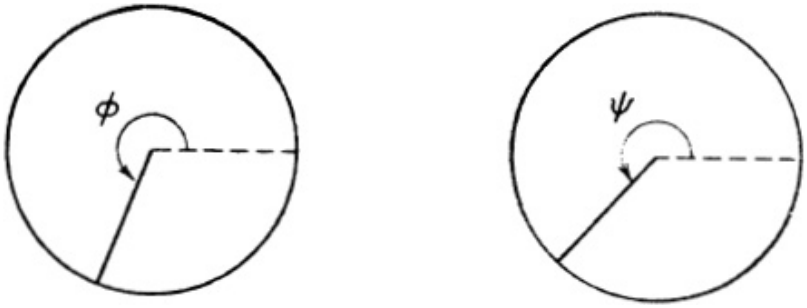


Figura 16.

Modifichiamo ora il dispositivo. Le aste sono ancora *OA* e *AB*, e *OA* ruota liberamente attorno a *O* in un piano fissato; ma *AB* è collegata a *OA* mediante uno snodo sferico in *A*, così che, fissata la posizione di *A*, può assumere qualsiasi direzione nello spazio. La configurazione è ora determinata da tre parametri: il primo è l'angolo ϕ che *OA* forma con il verso positivo dell'asse delle ascisse; gli altri due determinano la direzione spaziale di *AB*. Quest'ultima può essere individuata dal punto *B′* in cui il raggio *OB′*, parallelo e concorde ad *AB*, incontra la sfera unitaria con centro nell'origine *O*, oppure dalle due coordinate geografiche di *B′*. La varietà delle configurazioni è dunque tridimensionale; il lettore comprenderà facilmente che è il prodotto topologico di una circonferenza e di una sfera. È una varietà chiusa, compatta e priva di bordo, e pertanto non può essere realizzata come una figura contenuta nello spazio tridimensionale. Per farsene comunque un'idea intuitiva si può considerare la regione compresa fra due sfere concentriche. Ogni semiretta uscente dal centro comune le incontra in due punti: identificando ciascuna di queste coppie, ossia fondendola in un solo punto,

otteniamo una varietà tridimensionale che è il prodotto topologico di una sfera e di una circonferenza.

Possiamo complicare ancora il dispositivo: oltre a collegare *OA* e *AB* con uno snodo sferico in *A*, supponiamo che anche *OA* possa ruotare liberamente nello spazio attorno a *O*. L'insieme delle configurazioni diventa allora una varietà chiusa di quattro dimensioni, precisamente il prodotto di due sfere.

Abbiamo così visto che considerazioni meccaniche, o cinematiche, assai semplici conducono già a varietà topologiche di tre o più dimensioni. Ancora più importanti nello studio concreto e dettagliato dei problemi meccanici sono certe varietà, in generale multidimensionali, chiamate spazi delle fasi dei sistemi dinamici. Qui si tiene conto non soltanto delle configurazioni del sistema, ma anche delle velocità dei suoi punti. Limitiamoci a uno degli esempi più semplici: un punto che si muova liberamente su una circonferenza con velocità arbitraria. Ogni stato del sistema è determinato da due dati: la posizione del punto sulla circonferenza e la sua velocità, con segno, nell'istante considerato. La varietà degli stati, cioè lo spazio delle fasi, è evidentemente un cilindro infinito, prodotto di una circonferenza e di una retta.

La dimensione dello spazio delle fasi cresce con il numero dei gradi di libertà del sistema. Molte caratteristiche dinamiche di un sistema meccanico possono essere espresse mediante proprietà topologiche del suo spazio delle fasi. Per esempio, a ogni moto periodico non costante corrisponde una curva chiusa nello spazio delle fasi.

Lo studio degli spazi delle fasi dei sistemi dinamici che intervengono in meccanica, fisica e astronomia - meccanica celeste e cosmogonia - attirò l'attenzione dei matematici sulla topologia delle varietà multidimensionali. Fu proprio in rapporto a questi problemi che, negli anni Novanta dell'Ottocento, il celebre matematico francese Poincaré avviò la costruzione sistematica della topologia delle varietà, applicando il cosiddetto metodo combinatorio, che rimane uno dei metodi fondamentali della topologia.

§4. Il metodo combinatorio

Storicamente, il primo teorema della topologia è il teorema, o formula, di Eulero, che sembra fosse già noto a Cartesio. Esso afferma quanto segue. Prendiamo la superficie di un poliedro convesso arbitrario e indichiamo con α_0 il numero dei vertici, con α_1 quello degli spigoli e con α_2 quello delle facce. Vale allora la relazione seguente, nota come formula di Eulero:

$$\alpha_0 - \alpha_1 + \alpha_2 = 2. \qquad (1)$$

Questo teorema geometrico appartiene alla topologia perché la formula resta evidentemente valida quando sottoponiamo il poliedro convesso a una trasformazione topologica arbitraria. In generale gli spigoli cessano di essere rettilinei, le facce cessano di essere piane e la superficie del poliedro diventa curva; ma la relazione (1) fra i numeri di vertici, spigoli e facce, ora curvilinei, rimane valida. Il caso più importante è quello in cui tutte le facce sono triangoli: si ha allora una triangolazione, cioè una suddivisione della superficie in triangoli rettilinei o curvilinei. È facile ricondurvi il caso generale delle facce poligonali: basta suddividerle in triangoli, per esempio tracciando le diagonali da un vertice qualsiasi di ciascuna faccia. Possiamo dunque limitarci alle triangolazioni. Il metodo combinatorio nella topologia delle superfici consiste nel sostituire allo studio di una superficie quello di una sua triangolazione. Ci interessano naturalmente soltanto le proprietà indipendenti dalla scelta accidentale di una particolare triangolazione: essendo comuni a tutte le triangolazioni della superficie, esse esprimono una proprietà della superficie stessa.

La formula di Eulero conduce a una di queste proprietà, che esamineremo più attentamente. Il suo primo membro, $\alpha_0 - \alpha_1 + \alpha_2$, dove α_0, α_1 e α_2 sono rispettivamente i numeri di vertici, spigoli e triangoli, si chiama caratteristica di Eulero della triangolazione. Il teorema di Eulero afferma che, per ogni triangolazione di una superficie omeomorfa alla sfera, la caratteristica vale due. Ora, per qualsiasi superficie, non soltanto per quelle omeomorfe alla sfera, tutte le triangolazioni hanno la medesima caratteristica di Eulero.

È facile calcolare questa caratteristica per diverse superfici. Per la superficie laterale di un cilindro finito essa vale zero. Infatti, togliendo da una triangolazione della sfera gli interni di due triangoli non adiacenti e conservandone i bordi, otteniamo una superficie omeomorfa alla superficie laterale di un cilindro. Il numero di vertici e spigoli rimane invariato, mentre quello dei triangoli diminuisce di due: la caratteristica della nuova triangolazione è dunque zero. Togliamo ora dalla triangolazione di una sfera gli interni di $2p$ triangoli a due a due non adiacenti, cioè privi di vertici e lati comuni.[4] La caratteristica diminuisce di $2p$. È facile vedere che essa non cambia quando si attacca un tubo cilindrico a ciascuna coppia di fori: la caratteristica del tubo è zero e lungo i suoi bordi il numero di vertici uguaglia quello degli spigoli. Una superficie chiusa bilaterale di genere p ha quindi caratteristica di Eulero $2 - 2p$, risultato dimostrato per la prima volta dall'ammiraglio francese de Jonquières.

Introduciamo ora un'importante proprietà delle triangolazioni che soddisfa la condizione di invarianza topologica: se una triangolazione di una superficie la possiede, la possiedono tutte. È l'orientabilità. Osserviamo anzitutto che un triangolo può essere orientato, cioè dotato di un

verso di percorrenza del bordo. Ciascuna delle due orientazioni possibili è determinata da un ordine dei suoi vertici.[5] Consideriamo su una superficie due triangoli con un lato comune e nessun altro punto comune (figura 17). Le loro orientazioni si dicono compatibili se inducono versi opposti sul lato comune. Nel piano, o su qualsiasi superficie bilaterale, ciò significa che, guardandoli dalla stessa faccia, i due triangoli sono percorsi nello stesso senso: entrambi in senso antiorario oppure entrambi in senso orario. Una triangolazione di una superficie chiusa si dice orientabile se è possibile orientare tutti i triangoli in modo che due triangoli adiacenti lungo un lato abbiano sempre orientazioni compatibili. Vale allora il fatto seguente: ogni triangolazione di una superficie bilaterale è orientabile, mentre ogni triangolazione di una superficie unilaterale non lo è. Per questo le superfici bilaterali si chiamano anche orientabili e quelle unilaterali non orientabili. Scegliendo una triangolazione del nastro di Möbius, il lettore potrà facilmente convincersi della sua non orientabilità. Per ottenere una semplice suddivisione triangolare del piano proiettivo basta tracciarvi tre rette non passanti per uno stesso punto (figura 18). Esse lo dividono in quattro triangoli: uno è interamente nella parte finita del piano, mentre ciascuno degli altri tre è diviso in due parti dalla retta all'infinito. Nella figura 18 è tratteggiato uno dei triangoli che si estendono all'infinito. Dalla stessa figura si vede che ogni tentativo di orientare compatibilmente i quattro triangoli è destinato a fallire. In particolare, con le orientazioni indicate, la somma algebrica dei loro bordi non è zero, come dovrebbe essere per orientazioni compatibili, ma è la retta *AB* presa due volte.

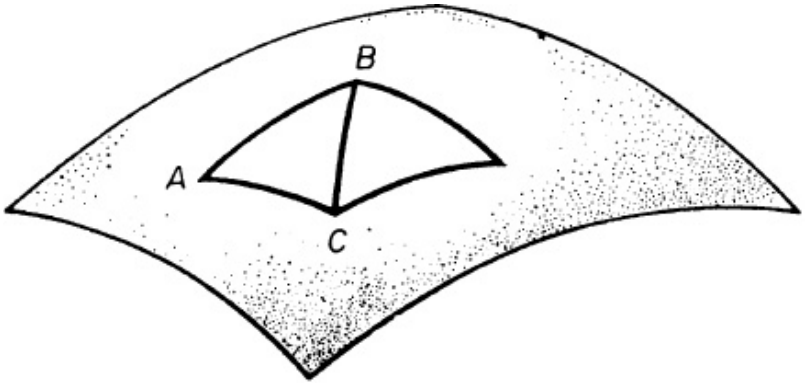


Figura 17.

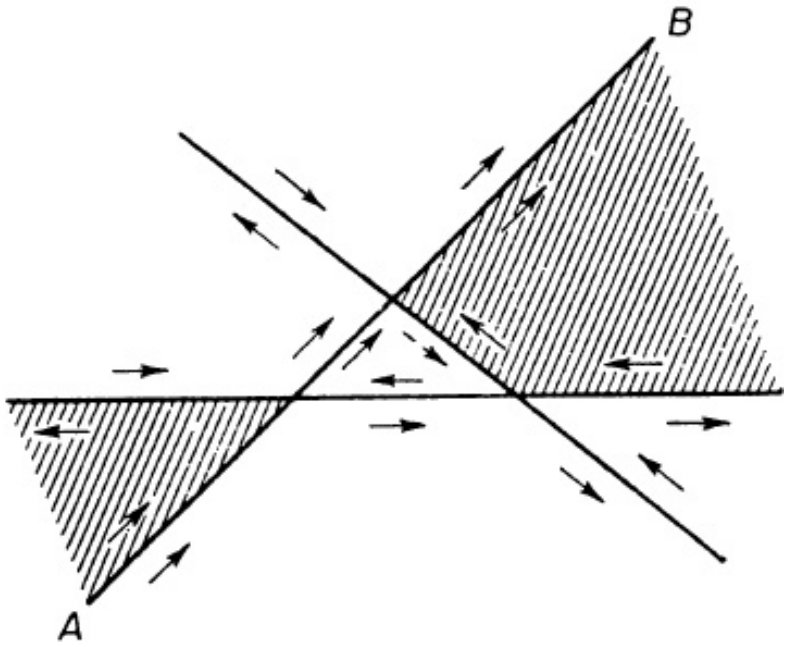


Figura 18.

La caratteristica di Eulero e l'orientabilità, o non orientabilità, costituiscono un sistema completo di invarianti topologici delle superfici chiuse connesse. Ciò significa che due di queste superfici sono omeomorfe se e solo se le loro triangolazioni hanno la stessa caratteristica di Eulero ed entrambe sono orientabili oppure entrambe non orientabili.

Il metodo combinatorio si applica non soltanto alle superfici, cioè alle varietà bidimensionali, ma anche alle varietà di dimensione arbitraria che ammettono una triangolazione. Nel caso tridimensionale, ai triangoli si sostituiscono i tetraedri: si parla di triangolazioni tridimensionali o suddivisioni simpliciali della varietà. La caratteristica di Eulero è allora $\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3$, dove α_i, per $i = 0, 1, 2, 3$, è il numero degli elementi di dimensione i: vertici, spigoli, facce bidimensionali e tetraedri. Per $n > 3$ si suddivide la varietà in simplessi n-dimensionali, ossia nei più semplici poliedri convessi di dimensione n, analoghi ai triangoli per $n = 2$ e ai tetraedri per $n = 3$. I simplessi e le loro facce costituiscono una triangolazione n-dimensionale. La caratteristica di Eulero si definisce ancora come la somma $\sum_{i=0}^{n}(-1)^i\alpha_i$, dove α_i è il numero degli elementi di dimensione i, per $i = 0, 1, 2, \ldots, n$. Essa ha lo stesso valore per tutte le triangolazioni della varietà e per quelle delle varietà a essa omeomorfe: è un invariante topologico. Allo stato delle conoscenze dell'epoca del presente testo, però, non si poteva ancora aspirare a un sistema completo di invarianti neppure per le varietà tridimensionali, nel senso in cui caratteristica di Eulero e orientabilità classificano le superfici.

Il metodo combinatorio ha grandissimo valore nella topologia contemporanea perché permette di applicare strumenti algebrici ai problemi topologici. Il lettore attento avrà già intravisto questa possibilità quando abbiamo parlato della somma algebrica dei bordi dei triangoli orientati in una suddivisione del piano proiettivo. Se un triangolo è orientato, cioè se è fissato il verso di percorrenza del suo contorno, è naturale considerarne il bordo come l'insieme dei lati, ciascuno dotato del verso indotto da tale percorrenza.

Consideriamo tutti i triangoli T_i^2, $i = 1, 2, \ldots, \alpha_2$, di una triangolazione della superficie. Ognuno ammette due orientazioni: indichiamo il triangolo T_i^2 con una delle orientazioni mediante t_i^2, e con l'orientazione opposta mediante $-t_i^2$. Allo stesso modo possiamo orientare ciascuno degli elementi unidimensionali, cioè gli spigoli T_k^1 $(k = 1, 2, \cdots, \alpha_1)$, assegnandogli uno dei due versi possibili. Indichiamo il segmento T_k^1 con un'orientazione mediante t_k^1, e con l'altra mediante $-t_k^1$. Se i lati del triangolo T_i^2 sono T_1^1, T_2^1, T_3^1, il bordo del triangolo orientato t_i^2 è costituito da questi stessi lati presi con versi determinati, ossia dai segmenti orientati $\epsilon_1 t_1^1$, $\epsilon_2 t_2^1$, $\epsilon_3 t_3^1$. Qui $\epsilon_i = 1$ se il verso indotto sullo spigolo T_i^1 coincide con il verso prescelto t_i^1, mentre $\epsilon_i = -1$ nel caso opposto. Il bordo di t_i^2 si indica con Δt_i^2. Come abbiamo visto, $\Delta t_i^2 = \sum \epsilon_k t_k^1$; possiamo pensare che la somma sia estesa a tutti gli spigoli della triangolazione, ponendo uguali a zero i coefficienti ϵ_k dei segmenti che non appartengono al bordo di t_i^2.

È naturale considerare allora somme più generali della forma $x^1 = \sum a_k t_k^1$, estese a tutti gli spigoli della triangolazione.[6] Il loro significato geometrico è semplice: ogni addendo rappresenta uno spigolo con un verso determinato e un coefficiente, cioè una «molteplicità». La somma algebrica rappresenta un percorso, o una combinazione di percorsi, formato da segmenti, ciascuno contato tante volte quante indica il coefficiente. Per esempio, percorriamo il poligono *ABCDEF* della figura 19 nel verso della freccia, passiamo lungo *AA′* al poligono *A′B′C′D′E′F′* e percorriamolo nel verso indicato; torniamo poi lungo *A′A* e percorriamo una seconda volta *ABCDEF* nello stesso verso. Nella somma ottenuta i segmenti *AB*, *BC*, *CD*, *DE*, *EF*, *FA* hanno coefficiente 2; i segmenti *A′B′*, *B′C′*, *C′D′*, *D′E′*, *E′F′*, *F′A′* hanno coefficiente 1; il segmento *AA′* non compare, cioè ha coefficiente zero, perché è percorso due volte in versi opposti.

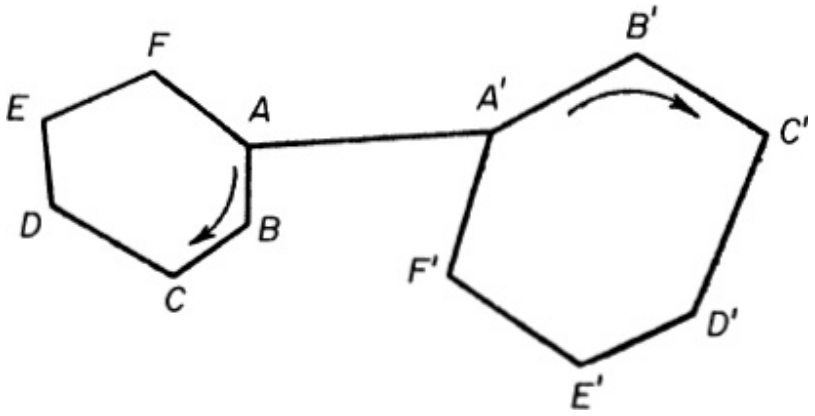


Figura 19.

Le somme della forma $x^1 = \sum a_k t_k^1$ si chiamano catene unidimensionali della triangolazione. Dal punto di vista algebrico sono forme lineari, cioè polinomi omogenei di primo grado; si possono sommare, sottrarre e moltiplicare per un intero arbitrario secondo le regole usuali dell'algebra. Fra esse hanno particolare importanza i cicli unidimensionali, che geometricamente corrispondono a percorsi chiusi, o a loro combinazioni. Nella figura 19 abbiamo appunto incontrato un percorso di questo tipo.

Per definire un ciclo in termini puramente algebrici, conveniamo che nel bordo del segmento orientato $\overrightarrow{AB}$ il vertice finale *B* compaia con il segno positivo, cioè con coefficiente +1, e il

vertice iniziale A con il segno negativo, cioè con coefficiente –1. Il bordo del segmento $\overrightarrow{AB}$ si scrive allora $\Delta(\overrightarrow{AB}) = B - A$.

Con questa convenzione osserviamo subito che la somma dei bordi dei segmenti $\overrightarrow{AB}$, $\overrightarrow{BC}$, $\overrightarrow{CD}$, ..., $\overrightarrow{FA}$, che formano un percorso chiuso nel senso usuale, è zero. Si giunge così alla definizione generale: un ciclo unidimensionale è una catena unidimensionale $z^1 = \Sigma_k a_k t_k^1$ per la quale la somma dei bordi degli addendi, cioè $\Sigma_k a_k \Delta t_k^1$, è nulla. È facile verificare che la somma di due cicli è ancora un ciclo. Anche moltiplicando un ciclo, come espressione algebrica, per un intero arbitrario si ottiene un ciclo. Possiamo quindi parlare di combinazioni lineari dei cicli $z_1^1, z_2^1, \cdots, z_s^1$, cioè di cicli della forma $z = \sum_{\nu=1}^{s} c_\nu z_\nu^1$, dove i c_ν sono interi.

Analogamente alle catene unidimensionali possiamo considerare le catene bidimensionali della triangolazione, cioè espressioni della forma $x^2 = \Sigma_i a_i t_i^2$, dove i t_i^2 sono triangoli orientati. Poiché il bordo di ogni triangolo orientato è un ciclo unidimensionale, anche la catena $\Sigma_i a_i \Delta t_i^2$ è un ciclo. Questo ciclo viene assunto come bordo Δx^2 della catena $x^2 = \Sigma_i a_i t_i^2$.

Il concetto di bordo permette ora di introdurre quello di omologia. Un ciclo unidimensionale z^1 si dice omologo a zero nella triangolazione se è il bordo di una sua catena bidimensionale. In qualsiasi triangolazione di una superficie convessa chiusa e, più in generale, di una superficie omeomorfa alla sfera, ogni ciclo unidimensionale è omologo a zero. Geometricamente il fatto è chiaro: ogni poligono semplice chiuso su una superficie convessa delimita una porzione della superficie. Sul toro non è così: né un meridiano né l'equatore costituiscono il bordo di una porzione della superficie. In ogni triangolazione del toro si possono trovare cicli analoghi al meridiano o all'equatore, e questi cicli non sono omologhi a zero.

Nella suddivisione del piano proiettivo che abbiamo costruito compare un fenomeno nuovo. Consideriamo come ciclo una retta, per esempio AB nella figura 18: essa non è omologa a zero. La stessa retta presa con coefficiente 2, invece, è omologa a zero. L'introduzione di coefficienti diversi da +1 nella definizione delle catene, che a prima vista sembra formale e superflua, permette dunque di riconoscere importanti proprietà geometriche delle superfici e, più in generale, delle varietà. Qui si tratta della torsione: esistono cicli non omologhi a zero nella varietà, che non sono bordi, ma che diventano omologhi a zero quando vengono moltiplicati per opportuni interi non nulli.

Introduciamo infine l'importante concetto di indipendenza omologica. I cicli $z_1, \ldots, z_s$ si dicono omologicamente indipendenti nella triangolazione se nessuna combinazione lineare $\sum c_i z_i$, con coefficienti interi non tutti nulli, è omologa a zero. Sul toro, un meridiano e un equatore, considerati come cicli di una triangolazione, forniscono un esempio di cicli omologicamente indipendenti.

Abbiamo definito i concetti fondamentali della topologia combinatoria - bordo, ciclo, omologia - per oggetti unidimensionali; essi si estendono però allo stesso modo a dimensioni arbitrarie. Per esempio, una catena bidimensionale $z^2 = \Sigma a_i t_i^2$ si chiama ciclo se il suo bordo

$\Delta z^2 = \sum a_i \Delta t_i^2$ è zero. Una catena tridimensionale è un'espressione della forma $x^3 = \sum a_h t_h^3$, dove i t_h^3 sono simplessi tridimensionali orientati, ossia tetraedri.

Come per il triangolo, l'orientazione di un simplesso tridimensionale, o tetraedro, è determinata da un ordine dei vertici; due ordini ottenibili l'uno dall'altro mediante una permutazione pari determinano la stessa orientazione. Il bordo del simplesso orientato t^3 = (*ABCD*) è la catena bidimensionale, che è anche un ciclo, Δt^3 = (*BCD*) − (*ACD*) + (*ABD*) − (*ABC*) (figura 20). Il bordo di una catena tridimensionale è la somma dei bordi dei suoi simplessi, ciascuno preso con lo stesso coefficiente con cui compare nella catena. Il lettore può verificare facilmente che il bordo di qualsiasi catena tridimensionale è un ciclo bidimensionale: basta dimostrarlo per un solo simplesso. Un ciclo bidimensionale è omologo a zero in una varietà se è il bordo di una catena tridimensionale della varietà, e così via. Dalla definizione di triangolazione orientabile o non orientabile[7] segue inoltre che, per una superficie chiusa connessa orientabile, esistono cicli bidimensionali non nulli, mentre per una superficie chiusa connessa non orientabile non ne esistono, considerando coefficienti interi. Questo risultato si estende alla dimensione massima delle varietà chiuse connesse di dimensione arbitraria.

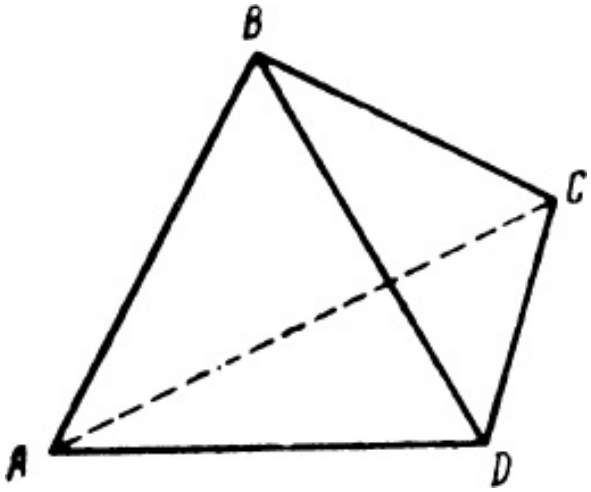


Figura 20.

I concetti introdotti permettono di definire l'ordine di connessione unidimensionale, bidimensionale e così via di una varietà di dimensione arbitraria. Il massimo numero di cicli omologicamente indipendenti di una data dimensione, presenti in una triangolazione della varietà, non dipende dalla scelta della triangolazione: si chiama ordine di connessione, o numero di Betti, della dimensione corrispondente.

Il numero di Betti unidimensionale di una superficie chiusa orientabile di genere *p* è 2*p*, cioè coincide con l'ordine di connessione definito nel §2. Quello del piano proiettivo è zero. Infatti ogni ciclo che non raggiunge l'infinito delimita, nel senso delle catene, una porzione del piano ed è quindi omologo a zero; un ciclo che si estende all'infinito, come una retta proiettiva, diventa omologo a zero quando è preso due volte. Il numero di Betti bidimensionale di ogni superficie chiusa non orientabile è zero: su una tale superficie non esiste alcun ciclo bidimensionale non nullo a coefficienti interi.

Il numero di Betti bidimensionale di una superficie chiusa connessa orientabile è invece uno. Orientando opportunamente tutti i triangoli di una sua triangolazione, otteniamo infatti un ciclo, detto ciclo fondamentale della superficie. È facile osservare che ogni ciclo bidimensionale è un multiplo intero di questo ciclo fondamentale. I risultati si estendono immediatamente alle varietà chiuse connesse *n*-dimensionali. Ricordiamo inoltre che il numero di Betti di dimensione zero di una varietà connessa, cioè non separabile in parti disgiunte e non in contatto, è uno.

I numeri di Betti delle diverse dimensioni sono collegati alla caratteristica di Eulero da una notevole formula dimostrata da Poincaré, che generalizza il teorema di Eulero. La formula di Eulero-Poincaré ha la semplice forma seguente:

$$\sum_{r=0}^{n} (-1)^r \alpha_r = \sum_{r=0}^{n} (-1)^r p_r .$$

A sinistra compare la caratteristica di Eulero di una triangolazione arbitraria della varietà; a destra, i numeri p_r sono i numeri di Betti delle diverse dimensioni *r*. Per una superficie chiusa connessa orientabile abbiamo visto che $p_0 = p_2 = 1$ e $p_1 = 2p$, dove *p* è il genere. Ritroviamo così il teorema di Eulero per queste superfici:

$$\alpha_0 - \alpha_1 + \alpha_2 = 2 - 2p.$$

§5. I campi vettoriali

Consideriamo la semplice equazione differenziale

$$\frac{dy}{dx} = F(x, y), \qquad (2)$$

in una regione piana G. Il suo significato geometrico è che in ogni punto (x, y) di G è assegnata una direzione di coefficiente angolare $F(x, y)$, dove F è una funzione continua. Diciamo che in G è dato un campo continuo di direzioni. Possiamo trasformarlo facilmente in un campo vettoriale continuo scegliendo, per esempio, in ciascuna direzione un vettore unitario con componente orizzontale positiva. Integrare l'equazione (2) significa suddividere, quando ciò è possibile, la regione in curve a due a due non intersecantisi, dette curve integrali, tali che in ogni punto la direzione assegnata sia quella della tangente all'unica curva integrale passante per quel punto.

Consideriamo, per esempio, l'equazione

$$\frac{dy}{dx} = \frac{y}{x}.$$

In ogni punto $M(x, y)$ diverso dall'origine, la direzione geometrica corrispondente è quella radiale $\overrightarrow{OM}$, dove O è l'origine. Le curve integrali sono contenute nelle rette passanti per O. Per ogni punto diverso da O passa un'unica curva integrale. L'origine è invece un punto singolare del campo radiale, un cosiddetto nodo, nel quale convergono tutte le rette integrali.

Se prendiamo l'equazione differenziale

$$\frac{dy}{dx} = -\frac{x}{y},$$

essa associa a ogni punto $M(x, y)$ diverso da O la direzione perpendicolare a $\overrightarrow{OM}$. In questo caso le curve integrali sono circonferenze con centro in O; l'origine è ancora un punto singolare, ma di tipo completamente diverso: è un centro, non un nodo. Esistono altri tipi di punti singolari (si veda il capitolo V, §6), alcuni dei quali sono rappresentati nelle figure 21 e 22. L'equazione $dy/dx = y/x$ non ha curve integrali chiuse; al contrario, il campo di direzioni descritto da $dy/dx = -x/y$ ha soltanto curve integrali chiuse. Sono possibili anche curve che si avvolgono a spirale attorno a un punto singolare, che in tal caso si chiama fuoco.

In molte applicazioni è particolarmente importante il caso del ciclo limite: una curva integrale chiusa isolata, attorno alla quale altre curve integrali si avvolgono a spirale. Sono possibili molte altre disposizioni reciproche delle curve integrali e dei punti singolari. I problemi riguardanti la forma e la disposizione delle curve integrali di un'equazione differenziale, nonché il numero, la natura e la posizione reciproca dei suoi punti singolari, appartengono alla teoria qualitativa delle equazioni differenziali. Come indica il nome, essa lascia da parte sia l'integrazione diretta «in forma finita» sia i metodi di approssimazione e di integrazione

numerica. Il suo oggetto fondamentale è, nella sostanza, la topologia del campo di direzioni e del sistema delle curve integrali dell'equazione.

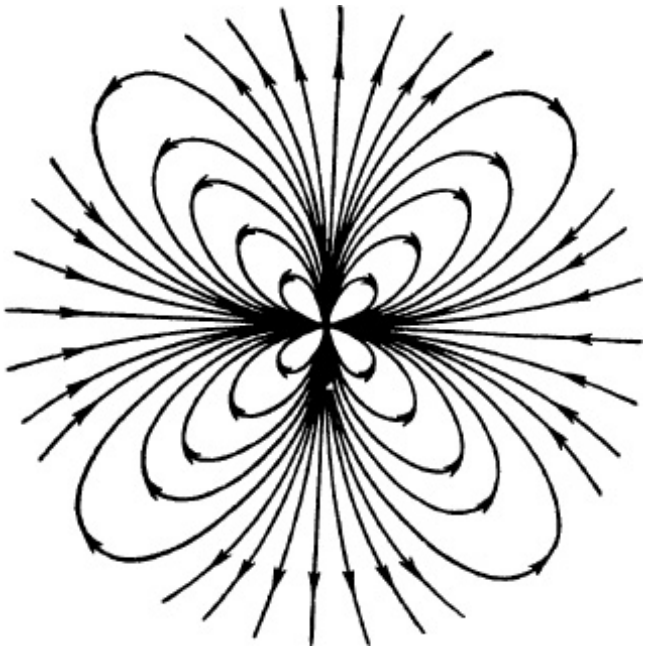


Figura 21.

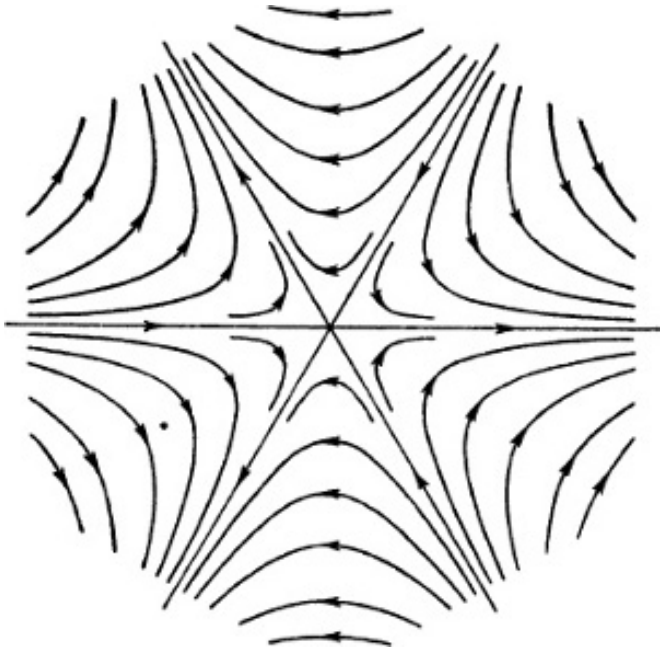


Figura 22.

Lo studio qualitativo delle equazioni differenziali, compresi i problemi sull'esistenza di curve integrali chiuse e, in particolare, sull'esistenza, il numero e la comparsa dei cicli limite, fu suggerito anzitutto da esigenze della meccanica, della fisica e della tecnica. Questi problemi emersero nelle ricerche di Poincaré in meccanica celeste e cosmogonia, che, come abbiamo ricordato, furono anche all'origine delle sue indagini topologiche. I problemi topologici delle equazioni differenziali occupano un posto centrale nelle importanti ricerche sovietiche sulla

teoria delle oscillazioni e sulla radiotecnica: ricordiamo la scuola di L. I. Mandel'štam e quella, da essa sviluppatasi, di A. A. Andronov, uno dei principali centri di ricerca sulla teoria qualitativa, ed essenzialmente topologica, delle equazioni differenziali. Un altro centro è la scuola di V. V. Stepanov e V. V. Nemyckij all'Università di Mosca. Anche nelle ricerche dei matematici di Leningrado, Sverdlovsk e Kazan' sulla teoria qualitativa i metodi topologici hanno un ruolo importante.

La teoria delle equazioni differenziali conduce a studiare campi vettoriali non soltanto nel piano, ma anche su varietà multidimensionali. Già i più semplici sistemi di più equazioni differenziali si interpretano geometricamente come campi di direzioni in spazi euclidei multidimensionali. Introdurre un integrale primo significa selezionare, fra tutte le curve integrali, quelle contenute in una determinata varietà di livello dell'integrale. Ogni sistema dinamico, nel senso classico, dà luogo in generale a una varietà multidimensionale dei suoi possibili stati (si veda il §3) e a un sistema di equazioni differenziali le cui curve integrali, distribuite nello spazio delle fasi, rappresentano i moti possibili. Ogni moto è determinato da opportune condizioni iniziali. L'oggetto fondamentale di studio è dunque il campo di direzioni e il sistema delle traiettorie sulla varietà. Le numerose applicazioni, soprattutto negli ultimi anni ai quali si riferisce il testo, spiegano perché la teoria qualitativa sia stata sviluppata nella massima generalità e perché si sia sviluppata, come suo fondamento, anche la topologia. Sono proprio questi temi meccanici, fisici e astronomici ad aver favorito la rapida crescita della topologia contemporanea, parte così importante dello sviluppo matematico della prima metà del Novecento. Il lettore che desideri conoscere i problemi topologici delle equazioni differenziali e i loro aspetti fisici e tecnici può consultare il noto libro di A. A. Andronov e S. E. Chajkin, *Theory of oscillations*, la cui edizione inglese uscì presso la Princeton University Press, Princeton, N. J., nel 1949.

Come esempio concreto della teoria dei campi vettoriali sulle varietà multidimensionali, consideriamo il problema del numero algebrico dei punti singolari di un campo.

Sia data una varietà liscia; per semplicità immaginiamo una superficie liscia chiusa. Supponiamo che in ogni punto sia assegnato un vettore tangente, dipendente con continuità dal punto sia in lunghezza sia in direzione, quando non è nullo. I punti singolari sono quelli in cui il vettore assegnato è nullo e non ha quindi una direzione determinata. Supponiamo che tali punti siano tutti isolati. Poiché la varietà è chiusa, e dunque compatta, essi sono in numero finito: altrimenti avrebbero un punto di accumulazione, che per continuità del campo sarebbe anch'esso singolare, ma non isolato.

A un punto singolare isolato possiamo attribuire un indice, concetto in un certo senso analogo alla molteplicità di una radice algebrica. Circondiamo il punto con una curva chiusa semplice C che lo isoli - nel piano basta una piccola circonferenza - e che non passi per punti singolari né ne racchiuda altri. In ogni punto di C la direzione del campo è univocamente determinata. Supponiamo per semplicità che un intorno del punto singolare comprendente C sia piano; nel caso generale possiamo riportare sul piano l'intorno e il campo mediante coordinate locali. Percorrendo la curva nel verso positivo, l'angolo formato dalla direzione del campo con una direzione fissa torna al valore iniziale aumentato di $2k\pi$, dove k è un intero ben determinato.

Questo numero si chiama indice del punto singolare, o numero di avvolgimento del campo lungo *C*. Esso non dipende dalla particolare curva scelta per isolare il punto. Le figure 21 e 22 mostrano punti singolari di indice rispettivamente +3 e −2. In modo analogo, ma più complesso, si definisce l'indice di un punto singolare di un campo vettoriale su una varietà di dimensione $n > 2$. Nel 1926 il matematico tedesco Hopf dimostrò il seguente notevole teorema: su una varietà liscia chiusa, la somma degli indici di un campo vettoriale continuo con un numero finito di punti singolari - cioè il numero algebrico di questi punti - non dipende dal campo ed è sempre uguale alla caratteristica di Eulero della varietà.

Dal teorema di Hopf segue che un campo vettoriale privo di singolarità può esistere soltanto su una varietà chiusa con caratteristica di Eulero zero. Su una varietà liscia chiusa connessa che soddisfi questa condizione è effettivamente sempre possibile costruirlo. Fra le superfici chiuse connesse, dunque, soltanto il toro e il cosiddetto toro unilaterale, ossia la bottiglia di Klein, ammettono campi vettoriali senza singolarità.

Alla teoria dei campi vettoriali è strettamente collegata quella delle applicazioni continue delle varietà in sé, in particolare i risultati sull'esistenza di punti fissi. Un punto *x* si chiama punto fisso di un'applicazione *f* se la sua immagine coincide con il punto stesso, cioè se

$$f(x) = x.$$

Per spiegare questo legame consideriamo il caso più semplice: un'applicazione continua *f* di un disco chiuso *K* in sé. Unendo ogni punto *x* di *K* alla sua immagine *f*(*x*), otteniamo il vettore $u_x = \overrightarrow{x, f(x)}$. Esso è nullo se e solo se *f*(*x*) = *x*, cioè se *x* è un punto fisso. Dimostriamo che un tale punto esiste. Supponiamo per assurdo che non esista e calcoliamo il numero di avvolgimento del campo lungo la circonferenza *C* che delimita *K*.

Sotto una deformazione continua del campo che non annulli i vettori su *C*, il numero di avvolgimento può variare soltanto con continuità; ma essendo intero deve rimanere costante. Ne segue che il numero di avvolgimento lungo *C* è 1. Infatti ogni punto del disco viene mandato nel disco stesso: perciò il vettore u_x di un punto *x* della circonferenza, che per ipotesi non è nullo, punta verso l'interno e forma un angolo acuto con il raggio diretto da *x* verso il centro *O*.

Deformiamo con continuità le direzioni di tutti i vettori u_x lungo *C*, ruotando ciascuno dell'angolo acuto necessario a farlo puntare verso il centro *O*. Il numero di avvolgimento non cambia. Alla fine il campo sulla circonferenza è diventato il campo radiale diretto verso il centro, il cui numero di avvolgimento è evidentemente uno. Anche il campo iniziale aveva quindi numero di avvolgimento uno.

Per continuità, i numeri di avvolgimento del campo lungo due circonferenze concentriche di raggi sufficientemente vicini sono uguali.[8] Il campo ha dunque numero di avvolgimento uno lungo tutte le circonferenze con centro *O* contenute in *K*. Ma per ipotesi il vettore u_x è definito e non nullo in tutti i punti del disco, compreso *O*; su una circonferenza sufficientemente piccola attorno a *O* il numero di avvolgimento deve allora essere zero. Abbiamo ottenuto una contraddizione: ogni applicazione continua di un disco chiuso in sé ha almeno un punto fisso.

È un caso particolare del fondamentale teorema di Brouwer: ogni applicazione continua di una palla chiusa *n*-dimensionale in sé possiede almeno un punto fisso.

Negli anni ai quali si riferisce il testo, il problema dell'esistenza di punti fissi per applicazioni di vario tipo è stato studiato approfonditamente ed è diventato una parte importante della topologia delle varietà.

§6. Lo sviluppo della topologia

La topologia delle superfici chiuse era l'unico settore della topologia già sviluppato in misura apprezzabile alla fine dell'Ottocento. La sua costruzione era legata alla teoria delle funzioni di variabile complessa, uno dei fenomeni più significativi della storia matematica di quel secolo, edificata attraverso metodi diversi. Fra i più fecondi per comprendere l'essenza dei fenomeni studiati vi era il metodo geometrico di Riemann. Esso mostrò in modo convincente che nella teoria generale delle funzioni complesse non era possibile limitarsi alle sole funzioni a un valore e condusse alla costruzione delle superfici di Riemann. Nel caso delle funzioni algebriche, queste superfici, opportunamente completate, sono superfici chiuse orientabili. Lo studio delle loro proprietà topologiche corrisponde, in un certo senso, allo studio della funzione algebrica considerata. Poincaré, Klein e i loro continuatori svilupparono le idee di Riemann, scoprendo legami profondi e inattesi fra la teoria delle funzioni, la topologia delle superfici chiuse e la geometria non euclidea, in particolare la teoria dei gruppi di movimenti del piano di Lobačevskij.[9] Fin dalle origini, dunque, la topologia si mostrò organicamente collegata a problemi fondamentali di settori assai diversi della matematica.

Proseguendo queste ricerche si vide che la sola topologia delle superfici non bastava: occorreva risolvere problemi di topologia *n*-dimensionale. Il primo fu quello dell'invarianza topologica della dimensione: dimostrare che non esiste un omeomorfismo di uno spazio euclideo *n*-dimensionale su uno spazio euclideo *m*-dimensionale quando $n \neq m$. Questo difficile problema fu risolto da Brouwer nel 1911.[10] La sua soluzione portò a nuovi metodi, grazie ai quali si svilupparono rapidamente i primi elementi della teoria delle applicazioni continue fra varietà multidimensionali e dei campi vettoriali su tali varietà. Tutte queste indagini richiedevano i concetti fondamentali della cosiddetta topologia insiemistica, nata dalla teoria generale degli insiemi costruita da Cantor nell'ultimo quarto dell'Ottocento.

Nella topologia insiemistica la classe delle figure geometriche studiate si amplia enormemente e comprende, se non tutti gli insiemi degli spazi euclidei, almeno tutti gli insiemi chiusi. Alla rapida crescita di questo nuovo indirizzo contribuirono studiosi di molti paesi; fra essi va ricordata soprattutto la scuola topologica polacca.

Le ricerche dei topologi sovietici aprirono una direzione sostanzialmente nuova. In particolare, P. S. Uryson (1898-1924), matematico di grande valore morto prematuramente, sviluppò la teoria generale della dimensione, fondando una classificazione di insiemi di punti molto generali secondo un criterio fondamentale: il numero delle dimensioni. Questa classificazione si rivelò estremamente feconda e introdusse punti di vista nuovi nello studio delle forme geometriche più generali.[11] Le idee di Uryson sulla dimensione furono un punto

di partenza per le importanti ricerche di L. A. Ljusternik, in collaborazione con L. G. Šnirel'man, sul calcolo delle variazioni.

Fra gli altri risultati, queste ricerche contengono una soluzione positiva completa di un celebre problema di Poincaré: l'esistenza di tre geodetiche chiuse semplici, cioè senza punti multipli, su ogni superficie riemanniana liscia omeomorfa alla sfera.

D'altra parte, sulla base della teoria della dimensione, P. S. Aleksandrov trasferì i metodi algebrici della topologia combinatoria nell'ambito insiemistico. Ne nacquero nuovi indirizzi di ricerca, nei quali i matematici sovietici, compresi quelli della generazione più giovane, occupavano un posto di primo piano.[12]

Per quanto riguarda la topologia combinatoria in senso stretto, dopo i lavori di Poincaré e Brouwer, attorno al 1915 iniziarono le ricerche dei topologi americani Veblen, Birkhoff, Alexander e Lefschetz, che ottennero risultati notevoli. Alexander dimostrò l'invarianza topologica dei numeri di Betti e un importante teorema di dualità, punto di partenza delle successive indagini di L. S. Pontrjagin. Lefschetz formulò una relazione per il numero algebrico dei punti fissi delle applicazioni continue delle varietà, ponendo le basi della teoria algebrica generale delle applicazioni continue, poi sviluppata da Hopf. A Birkhoff si devono progressi essenziali nella teoria dei sistemi dinamici, sia sul piano topologico sia su quello metrico. Hopf ottenne ulteriori risultati profondi sulle varietà e sulle loro applicazioni continue: fra l'altro dimostrò l'esistenza di infinite applicazioni continue della sfera tridimensionale sulla sfera bidimensionale, essenzialmente differenti nel senso che nessuna coppia può essere collegata da una deformazione continua. Divenne così il fondatore di un nuovo indirizzo, la teoria dell'omotopia. Più recentemente, rispetto all'epoca del testo, un forte impulso alla teoria dell'omotopia e all'intera topologia combinatoria è venuto dalla nuova scuola francese, con Leray, Serre e altri.

Come abbiamo ricordato, le ricerche fondamentali di Uryson segnarono l'inizio dell'attività dei matematici sovietici in topologia. Esse riguardavano la topologia insiemistica, ma già alla fine degli anni Venti l'interesse si estese anche a quella combinatoria. Ciò avvenne in modo originale, applicando metodi combinatori allo studio degli insiemi chiusi, oggetti di natura molto generale. Su questa base nacque una delle più importanti scoperte geometriche del Novecento: la formulazione e la dimostrazione, da parte di L. S. Pontrjagin, della sua legge generale di dualità, che stabilisce legami profondi e in un certo senso esaurienti fra la struttura topologica di un insieme chiuso nello spazio euclideo n-dimensionale e le parti del suo complementare. In relazione a questa legge, Pontrjagin costruì una teoria generale dei caratteri dei gruppi commutativi, che lo condusse a ulteriori indagini sui gruppi topologici generali e sui gruppi continui classici di Lie, settore profondamente trasformato dai suoi lavori. Successivamente Pontrjagin e i suoi allievi, fra cui V. G. Boltjanskij e M. M. Postnikov, svolsero importanti ricerche sulla topologia delle varietà e sulle loro applicazioni continue. Vi applicarono un nuovo metodo, la cosiddetta ∇-omologia, o coomologia, introdotta nella topologia combinatoria da A. N. Kolmogorov e, indipendentemente, da Alexander. Questo metodo, divenuto centrale nella teoria dell'omotopia, permise di sviluppare la dualità di Pontrjagin in direzioni diverse, conducendo ai teoremi di dualità di A. N. Kolmogorov e

Alexander, di P. S. Aleksandrov e di K. A. Sitnikov, annoverati fra i risultati più significativi della topologia dell'epoca. Il metodo trovò inoltre importanti applicazioni nelle ricerche allora recenti di L. A. Ljusternik sul calcolo delle variazioni.

§7. Spazi metrici e topologici

All'inizio della nostra esposizione abbiamo presentato il contatto fra parti di una figura come concetto topologico fondamentale e definito continue le trasformazioni che conservano tale relazione. Non ne abbiamo però ancora dato una definizione rigorosa. Per farlo con sufficiente generalità occorrono i concetti della teoria degli insiemi. È questo il compito dell'ultima sezione, che porterà al concetto di spazio topologico.

La teoria degli insiemi ha permesso di dare al concetto di figura geometrica un'ampiezza e una generalità irraggiungibili nella matematica cosiddetta classica. Oggetto dello studio geometrico, e in particolare topologico, diventa un insieme arbitrario di punti, cioè qualsiasi insieme i cui elementi siano punti di uno spazio euclideo n-dimensionale. Fra questi punti è definita una distanza: fra $A = (x_1, x_2, \ldots, x_n)$ e $B = (y_1, y_2, \ldots, y_n)$ essa è, per definizione, il numero non negativo

$$\rho(A, B) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

La distanza permette di definire il contatto dapprima fra un punto e un insieme e poi fra due insiemi. Un punto A si dice aderente all'insieme M se M contiene punti la cui distanza da A è minore di qualsiasi numero positivo prefissato. Ogni punto di un insieme è evidentemente aderente a esso; ma possono esservi anche punti aderenti che non gli appartengono. Prendiamo, per esempio, l'intervallo aperto (0, 1) della retta numerica, costituito dai punti compresi fra 0 e 1. I punti 0 e 1 non appartengono all'intervallo, ma gli sono aderenti, perché esso contiene punti arbitrariamente vicini a ciascuno dei due estremi. Un insieme si dice chiuso se contiene tutti i propri punti aderenti. È chiuso, per esempio, l'intervallo [0, 1], costituito dai punti x che soddisfano $0 \leqslant x \leqslant 1$. Gli insiemi chiusi del piano e, ancor più, degli spazi di tre o più dimensioni possono avere una struttura estremamente complicata: costituiscono infatti uno degli oggetti principali della topologia insiemistica dello spazio n-dimensionale.

Diciamo poi che due insiemi P e Q sono in contatto se almeno uno contiene punti aderenti all'altro. Ne segue che due insiemi chiusi possono essere in contatto soltanto se hanno almeno un punto comune. Gli intervalli [0, 1] e (1, 2), invece, non hanno punti comuni, ma sono in contatto perché 1 appartiene al primo ed è aderente al secondo. Possiamo ora dire che un insieme S contenuto in R separa R, oppure che è una sua «sezione», se il complementare $R - S$, formato dai punti di R non appartenenti a S, si può rappresentare come unione di due insiemi non vuoti che non sono in contatto.

Le idee di Lobačevskij sul contatto e sulla separazione ricevono così una formulazione rigorosa e molto generale nella topologia contemporanea. Abbiamo visto che su di esse si fonda la definizione di dimensione di Uryson (si veda la nota al §6), che può ora essere

precisata rigorosamente. Lo stesso vale per le applicazioni continue: un'applicazione *f* da un insieme *X* a un insieme *Y* si dice continua se conserva l'aderenza, cioè se, per ogni sottoinsieme *P* di *X*, il fatto che un punto *A* di *X* sia aderente a *P* implica che *f*(*A*) sia aderente a *f*(*P*). Infine una corrispondenza biunivoca di *X* su *Y* si dice topologica se è continua e se è continua anche la sua inversa da *Y* a *X*. Queste definizioni danno un fondamento preciso a quanto abbiamo esposto nelle prime sezioni.

La topologia insiemistica non si arresta però alla generalità raggiunta considerando come figure geometriche tutti gli insiemi di punti. È naturale introdurre la distanza non soltanto fra i punti degli spazi euclidei, ma anche fra oggetti che sembrano non avere alcun rapporto con la geometria.

Consideriamo, per esempio, l'insieme di tutte le funzioni continue sull'intervallo [0, 1]. Possiamo definire la distanza $\rho(f, g)$ fra due funzioni come il massimo di $|f(x) - g(x)|$ al variare di x nell'intero intervallo. Questa distanza possiede tutte le proprietà fondamentali della distanza fra punti: è zero se e solo se le due funzioni coincidono, cioè se $f(x) = g(x)$ per ogni x; è simmetrica, ossia $\rho(f, g) = \rho(g, f)$; infine soddisfa la disuguaglianza triangolare: per tre funzioni qualsiasi f_1, f_2, f_3 vale $\rho(f_1, f_2) + \rho(f_2, f_3) \geqslant \rho(f_1, f_3)$. Si dice che questa distanza rende l'insieme delle funzioni uno spazio metrico, solitamente indicato con *C*. Più in generale, uno spazio metrico è un insieme di oggetti arbitrari, detti punti dello spazio, fra ogni coppia dei quali è definita una distanza: un numero non negativo che soddisfa gli assiomi appena enunciati.

In uno spazio metrico arbitrario possiamo allora parlare di punti aderenti ai suoi sottoinsiemi, di contatto fra sottoinsiemi e, in generale, dei concetti topologici: insiemi chiusi, applicazioni continue e altri concetti costruiti su questi. Si apre così un campo vasto e fecondo di applicazioni delle idee topologiche e geometriche a oggetti matematici per i quali sembrerebbe impossibile parlare di geometria. Vediamo un esempio.

Riprendiamo l'equazione differenziale (2):

$$\frac{dy}{dx} = F(x, y).$$

Se $y = \phi(x)$, per $0 \leqslant x \leqslant 1$, è una soluzione che assume per esempio il valore $y = 0$ in $x = 0$, allora ϕ soddisfa evidentemente l'equazione integrale

$$\phi(x) = \int_0^x F[t, \phi(t)]\, dt. \quad (3)$$

Consideriamo ora l'integrale $G(f)(x) = \int_0^x F[t, f(t)]\, dt$, dove $0 \leqslant x \leqslant 1$ e *f* è una funzione continua arbitraria su [0, 1]. Supponendo che *F* sia definita e continua nei punti necessari, questo integrale è una funzione continua $g(x)$ sullo stesso intervallo. L'espressione $G(f)(x) = \int_0^x F[t, f(t)]\, dt$ associa dunque a ogni funzione *f* una funzione $g = G(f)$: abbiamo un'applicazione *G* dello spazio metrico *C* in sé, la cui continuità si verifica facilmente. Come si caratterizza una funzione $\phi(x)$, eventualmente non unica, che risolve l'equazione (2) con il dato iniziale e quindi l'equazione equivalente (3)? Essa viene trasformata in sé stessa: è un punto fisso di *G*. L'esistenza di un tale punto fisso può essere dedotta, sotto le opportune

ipotesi e scegliendo un intervallo e un insieme di funzioni adatti, da un teorema generale sulle applicazioni continue negli spazi metrici dimostrato nel 1926 da P. S. Aleksandrov e V. V. Nemyckij. La sola continuità non garantisce invece una soluzione sull'intero intervallo prefissato [0, 1]. Lo studio di spazi metrici i cui punti sono funzioni, detti spazi funzionali, è ormai uno strumento abituale dell'analisi. Lo studio degli spazi funzionali con metodi in parte topologici, ma soprattutto algebrici in senso ampio, costituisce il contenuto dell'analisi funzionale (capitolo XIX).

Come ricordato nel capitolo introduttivo, l'analisi funzionale occupa un posto di grande rilievo nella matematica contemporanea, per la varietà dei legami con altri settori e per il suo valore nelle scienze naturali, soprattutto nella fisica teorica. Lo studio delle proprietà topologiche degli spazi funzionali è strettamente connesso al calcolo delle variazioni e alla teoria delle equazioni alle derivate parziali, come mostrano le ricerche di Ljusternik, Morse, Leray, Schauder, Krasnosel'skij e altri. In esse hanno un ruolo importante i problemi sull'esistenza di punti fissi per applicazioni continue degli spazi funzionali.

La topologia degli spazi funzionali e, più in generale, degli spazi metrici non rappresenta il massimo grado di generalità delle teorie topologiche contemporanee. Negli spazi metrici, infatti, il concetto fondamentale di aderenza è introdotto mediante la distanza, che non è a sua volta un concetto topologico. Nasce quindi il problema di definire l'aderenza direttamente, per via assiomatica. Si giunge così al concetto di spazio topologico, il più generale fra quelli qui considerati.

Uno spazio topologico è un insieme di oggetti di natura arbitraria, chiamati punti dello spazio, nel quale per ogni sottoinsieme sono assegnati, in qualche modo, i suoi punti aderenti. Questa assegnazione deve soddisfare alcune condizioni naturali, dette assiomi dello spazio topologico: per esempio, ogni punto di un insieme gli è aderente; un punto aderente all'unione di due insiemi è aderente ad almeno uno dei due; e così via. La teoria degli spazi topologici è oggetto di ricerche profonde, al cui sviluppo hanno contribuito in modo rilevante i matematici sovietici P. S. Uryson, P. S. Aleksandrov, A. N. Tichonov e altri. Fra i risultati allora più recenti va ricordato quello fondamentale del giovane matematico Ju. M. Smirnov, che trovò condizioni necessarie e sufficienti per la metrizzabilità di uno spazio topologico: condizioni che permettono di definire fra i punti una distanza tale da indurre esattamente la topologia assegnata. In altre parole, i punti aderenti di ogni sottoinsieme nello spazio metrico così ottenuto devono coincidere con quelli definiti inizialmente nello spazio topologico.

Letture consigliate

P. S. Aleksandrov, Combinatorial topology, 3 vols., Graylock Press, Albany, N. Y., vol. 1, 1956; vol. 2, 1957; vol. 3, 1960.

S. S. Cairns, Introductory topology, Ronald Press, New York, 1961.

D. W. Hall and G. L. Spencer, Elementary topology, Wiley, New York, 1955.

P. J. Hilton and S. Wiley, Homology theory: an introduction to algebraic topology, Cambridge University Press, New York, 1960.

J. G. Hocking and G. S. Young, Topology, Addison-Wesley, Reading, Mass., 1961.

J. L. Kelley, General topology, Van Nostrand, New York, 1955.

S. Lefschetz, Introduction to topology, Princeton University Press, Princeton, N. J., 1949.

M. H. A. Newman, Elements of the topology of plane sets of points, 2nd ed., Cambridge University Press, New York, 1951.

L. S. Pontrjagin, Foundations of combinatorial topology, Graylock Press, Albany, N. Y., 1952.

Note al capitolo XVIII

[1] N. I. Lobačevskij, *Opere complete*, vol. II, Gostehizdat, 1949, p. 168.

[2] Commenti ai *Nuovi principi della geometria*, nell'opera citata, p. 465.

[3] Si veda, per esempio, il capitolo XVII, §6, oppure il libro di P. S. Aleksandrov *Che cos'è la geometria non euclidea*, Mosca, 1951.

[4] Basta rendere la triangolazione sufficientemente fine, cosa sempre possibile mediante un'opportuna suddivisione di una triangolazione qualsiasi.

[5] È facile vedere che due ordini dei vertici determinano la stessa orientazione, cioè lo stesso verso di percorrenza, se e solo se differiscono per una permutazione pari. Così (*ABC*), (*BCA*), (*CAB*) determinano un'orientazione del triangolo, mentre (*BAC*), (*ACB*), (*CBA*) determinano l'altra. Per le permutazioni pari e dispari si veda, per esempio, il capitolo XX, §3.

[6] I coefficienti a_k si suppongono interi.

[7] La definizione, data per le triangolazioni delle superfici, si estende alle triangolazioni delle varietà di dimensione arbitraria.

[8] Poiché si suppone che *f* non abbia punti fissi, il campo è definito e non nullo ovunque; possiamo quindi considerarne il numero di avvolgimento lungo qualsiasi curva chiusa contenuta in *K*.

[9] A questo proposito si veda A. I. Markuševič, *Teoria delle funzioni analitiche*, Gostehizdat, 1950.

[10] Lo sviluppo della teoria delle funzioni richiedeva in realtà la soluzione di un problema ancora più difficile: dimostrare che l'immagine topologica di una regione aperta *n*-dimensionale, contenuta nello spazio *n*-dimensionale, è ancora una regione aperta quando l'immagine è nello stesso spazio. Anche questo problema, noto come invarianza del dominio, fu risolto da Brouwer.

[11] La definizione induttiva di dimensione proposta da Uryson può essere considerata uno sviluppo particolarmente completo dell'idea di Lobačevskij di assumere la separazione come operazione geometrica fondamentale. In termini approssimativi, un insieme è zero-dimensionale se può essere decomposto in parti arbitrariamente piccole, a due a due non in contatto. È *n*-dimensionale se può essere «separato» mediante sottoinsiemi di dimensione $n - 1$ in parti arbitrariamente piccole, a due a due non in contatto, e se non è possibile farlo con sottoinsiemi di dimensione inferiore a $n - 1$. È una descrizione intuitiva, non la definizione formale completa. Il significato preciso di contatto nella topologia contemporanea è dato nel §7.

[12] Vanno ricordate la teoria omologica della dimensione di P. S. Aleksandrov, le notevoli costruzioni a essa collegate di L. S. Pontrjagin e gli sviluppi successivi nelle ricerche di M. S. Bokštejn, V. G. Boltjanskij e, in particolare, K. A. Sitnikov. La legge di dualità di L. S. Pontrjagin è discussa più avanti nel testo principale.

Nota editoriale

Sono conservate le sette sezioni, le 22 figure numerate, tutte le formule, la bibliografia e le 12 note. La figura 11, troncata nella copia elettronica, è recuperata integralmente dal DjVu. Nelle formule integrali del §7 si distingue la variabile d'integrazione t dall'estremo x. Sono corretti alcuni refusi nelle iniziali dei nomi, nei vertici del poligono della figura 19, nel simbolo della distanza e nella definizione di continuità: il sottoinsieme P appartiene a X, non a Y. Gli indici delle singolarità nelle figure 21 e 22 sono rispettivamente +3 e −2; nel testo inglese risultano scambiati.

«Superficie chiusa» e «varietà chiusa» indicano qui oggetti compatti e senza bordo; nei teoremi di classificazione e nei risultati sul ciclo fondamentale è sottintesa, ed è resa esplicita, anche la connessione. Le curve che separano la sfera sono curve chiuse semplici. L'ordine di connessione del §2 è quello storico adottato dall'autore, pari a 2p sulle superfici orientabili di genere p: le curve di taglio possono incontrarsi, come il meridiano e il parallelo del toro, e si intendono sistemi finiti di curve semplici in posizione regolare. Non va confuso con il massimo numero di tagli lungo curve disgiunte che lasciano connessa la superficie.

La suddivisione del piano proiettivo della figura 18 è una decomposizione triangolare con identificazioni; non è un complesso simpliciale astratto in senso stretto, ma può essere ulteriormente suddivisa. Le catene hanno coefficienti interi; i risultati sui cicli di dimensione massima riguardano varietà chiuse connesse. Il discorso sulle triangolazioni in dimensione arbitraria si riferisce alle varietà triangolabili e non afferma che ogni varietà topologica ammetta una triangolazione.

Nel teorema del punto fisso si tratta del disco chiuso e della palla chiusa, non della sola circonferenza o superficie sferica. Per i campi esemplificati da dy/dx = y/x e dy/dx = −x/y si considera il prolungamento geometrico delle direzioni anche dove il rapporto non è definito; le traiettorie non attraversano l'equilibrio in tempo finito come soluzioni di un sistema regolare. La sola continuità del termine di un'equazione differenziale non garantisce né unicità delle soluzioni né esistenza sull'intero intervallo [0, 1]: nel §7 sono richiamate le ipotesi necessarie all'applicazione dei teoremi di punto fisso.

La descrizione iniziale delle trasformazioni come deformazioni della gomma è intuitiva: la definizione precisa di omeomorfismo è quella del §7 e non richiede una deformazione dell'ambiente. La ricostruzione storica, le valutazioni sui diversi indirizzi e i riferimenti a risultati recenti rispecchiano l'epoca dell'opera; non costituiscono un resoconto dello stato attuale della topologia.

CAPITOLO XIX

Analisi funzionale

I. M. Gel'fand

La nascita e la diffusione dell'analisi funzionale nel Novecento ebbero due cause principali. Da un lato si avvertì l'esigenza di interpretare da un punto di vista unitario il vasto materiale accumulato nell'Ottocento in settori matematici diversi, spesso scarsamente collegati. I concetti fondamentali dell'analisi funzionale si formarono e si precisarono sotto aspetti differenti e per ragioni diverse. Molti emersero naturalmente nello sviluppo del calcolo delle variazioni, nei problemi sulle oscillazioni - passando da sistemi con un numero finito di gradi di libertà ai mezzi continui -, nella teoria delle equazioni integrali, nelle equazioni differenziali ordinarie e alle derivate parziali, con i problemi al contorno e agli autovalori, nella teoria delle funzioni di variabile reale, nel calcolo operatoriale, nella teoria dell'approssimazione delle funzioni e in altri campi. L'analisi funzionale permise di comprendere molti risultati da un unico punto di vista e spesso favorì la scoperta di risultati nuovi. Nei decenni successivi i concetti e gli strumenti così preparati furono impiegati in un nuovo settore della fisica teorica: la meccanica quantistica.

D'altro lato, lo studio dei problemi matematici legati alla meccanica quantistica divenne decisivo per lo sviluppo dell'analisi funzionale stessa, dando origine a suoi indirizzi fondamentali e continuando ad alimentarne la crescita.

L'analisi funzionale è ancora ben lontana dall'essere compiuta. Al contrario, nel suo ulteriore sviluppo i problemi e le esigenze della fisica contemporanea avranno senza dubbio, secondo la prospettiva dell'autore, un'importanza paragonabile a quella che la meccanica classica ebbe per la nascita e lo sviluppo del calcolo differenziale e integrale nel Settecento.

Non è possibile includere in questo capitolo tutti i problemi dell'analisi funzionale, e neppure tutti quelli fondamentali: molti settori importanti oltrepassano i limiti del libro. Limitandoci ad alcuni problemi scelti, vogliamo tuttavia presentare al lettore i concetti essenziali e illustrare, per quanto possibile, i collegamenti appena descritti. I problemi qui trattati furono studiati soprattutto all'inizio del Novecento, a partire dai lavori classici di Hilbert, uno dei fondatori dell'analisi funzionale. Da allora la disciplina si è sviluppata vigorosamente e ha trovato ampie applicazioni in quasi ogni ramo della matematica: equazioni alle derivate parziali, probabilità, meccanica quantistica, teoria quantistica dei campi e così via. Purtroppo non possiamo esporre questi sviluppi successivi: occorrerebbe un altro ampio volume. Ci limiteremo perciò a uno dei problemi più antichi, la teoria delle autofunzioni.

§1. Lo spazio a n dimensioni

Nel seguito useremo i concetti fondamentali dello spazio a n dimensioni. Benché siano già stati introdotti nei capitoli sull'algebra lineare e sugli spazi astratti, riteniamo utile riprenderli nella forma in cui serviranno qui. Per leggere questa sezione bastano i fondamenti della geometria analitica.

Sappiamo che, nella geometria analitica dello spazio tridimensionale, un punto è individuato da una terna di numeri (f_1, f_2, f_3), le sue coordinate. La distanza del punto dall'origine è $\sqrt{f_1^2+f_2^2+f_3^2}$. Se consideriamo il punto come estremo del vettore che parte dall'origine, anche la lunghezza del vettore è $\sqrt{f_1^2+f_2^2+f_3^2}$. Il coseno dell'angolo fra i vettori non nulli diretti dall'origine ai punti $A(f_1,f_2,f_3)$ e $B(g_1, g_2, g_3)$ è dato dalla formula

$$\cos\phi = \frac{f_1g_1 + f_2g_2 + f_3g_3}{\sqrt{f_1^2+f_2^2+f_3^2}\,\sqrt{g_1^2+g_2^2+g_3^2}}.$$

Dalla trigonometria sappiamo che $|\cos\phi| \leqslant 1$. Abbiamo quindi la disuguaglianza

$$\frac{|f_1g_1 + f_2g_2 + f_3g_3|}{\sqrt{f_1^2+f_2^2+f_3^2}\,\sqrt{g_1^2+g_2^2+g_3^2}} \leqslant 1,$$

e di conseguenza vale sempre

$$(f_1g_1 + f_2g_2 + f_3g_3)^2 \leqslant (f_1^2+f_2^2+f_3^2)(g_1^2+g_2^2+g_3^2). \qquad (1)$$

Quest'ultima disuguaglianza ha carattere algebrico ed è vera per sei numeri reali arbitrari (f_1, f_2, f_3) e (g_1, g_2, g_3), perché essi possono essere le coordinate di due punti dello spazio. Eppure la (1) è stata ottenuta da considerazioni puramente geometriche: il suo stretto legame con la geometria le conferisce un significato intuitivo.

Quando si esprimono analiticamente relazioni geometriche, spesso i risultati rimangono validi sostituendo a una terna di numeri una successione di n numeri. Per esempio, la disuguaglianza (1) si generalizza a $2n$ numeri reali $(f_1, f_2, \ldots, f_n)$ e $(g_1, g_2, \ldots, g_n)$: per qualsiasi scelta di questi numeri vale la disuguaglianza analoga

$$(f_1g_1 + f_2g_2 + \cdots + f_ng_n)^2 \leqslant (f_1^2+f_2^2+\cdots+f_n^2)(g_1^2+g_2^2+\cdots+g_n^2). \qquad (1')$$

Questa disuguaglianza, di cui la (1) è un caso particolare, può essere dimostrata per via puramente analitica.[1] Molte altre relazioni fra terne di numeri ricavate in geometria analitica si generalizzano allo stesso modo. Il rapporto fra geometria e relazioni numeriche diventa particolarmente chiaro introducendo lo spazio a n dimensioni, già considerato nel capitolo XVI. Ne richiamiamo brevemente la definizione.

Una successione di n numeri reali $(f_1, f_2, \ldots, f_n)$ si chiama punto o vettore dello spazio a n dimensioni; useremo più spesso la seconda denominazione. D'ora in poi indicheremo questo vettore con la sola lettera f.

Come nello spazio tridimensionale la somma dei vettori si ottiene sommando le componenti, così definiamo la somma dei vettori

$$f = \{f_1, f_2, \ldots, f_n\} \text{ e } g = \{g_1, g_2, \ldots, g_n\}.$$

come il vettore $\{f_1 + g_1, f_2 + g_2, \ldots, f_n + g_n\}$, che indichiamo con $f + g$.

Il prodotto del vettore $f = \{f_1, f_2, \ldots, f_n\}$ per il numero λ è il vettore $\lambda f = \{\lambda f_1, \lambda f_2, \ldots, \lambda f_n\}$.

La lunghezza del vettore $f = \{f_1, f_2, \ldots, f_n\}$ si definisce, come nello spazio tridimensionale, mediante $\sqrt{f_1^2 + f_2^2 + \cdots + f_n^2}$.

Anche l'angolo ϕ fra due vettori non nulli $f = \{f_1, f_2, \ldots, f_n\}$ e $g = \{g_1, g_2, \ldots, g_n\}$ è definito dal suo coseno, esattamente come in tre dimensioni, attraverso la formula[2]

$$\cos\phi = \frac{f_1 g_1 + f_2 g_2 + \cdots + f_n g_n}{\sqrt{f_1^2 + f_2^2 + \cdots + f_n^2}\,\sqrt{g_1^2 + g_2^2 + \cdots + g_n^2}}. \qquad (2)$$

Il prodotto scalare di due vettori è il prodotto delle loro lunghezze per il coseno dell'angolo compreso. Poiché le lunghezze di $f = \{f_1, \ldots, f_n\}$ e $g = \{g_1, \ldots, g_n\}$ sono rispettivamente $\sqrt{f_1^2 + f_2^2 + \cdots + f_n^2}$ e $\sqrt{g_1^2 + g_2^2 + \cdots + g_n^2}$, il loro prodotto scalare, indicato con (f, g), è dato da

$$(f, g) = f_1 g_1 + f_2 g_2 + \cdots + f_n g_n. \qquad (3)$$

In particolare, per vettori non nulli la condizione di ortogonalità, cioè di perpendicolarità, è $\cos\phi = 0$, ossia $(f, g) = 0$. Quest'ultima uguaglianza definisce l'ortogonalità anche quando uno dei vettori è nullo.

Dalla formula (3) il lettore può verificare le seguenti proprietà del prodotto scalare:

1. $(f, g) = (g, f)$.

2. $(\lambda f, g) = \lambda(f, g)$.

3. $(f, g_1 + g_2) = (f, g_1) + (f, g_2)$.

4. $(f, f) \geqslant 0$, con uguaglianza soltanto per $f = 0$, cioè quando $f_1 = f_2 = \ldots = f_n = 0$.

Il prodotto scalare di un vettore con sé stesso, (f, f), è uguale al quadrato della sua lunghezza.

Il prodotto scalare è uno strumento molto comodo nello studio degli spazi a n dimensioni. Non ne svilupperemo qui la geometria, ma ci limiteremo a un esempio.

Scegliamo il teorema di Pitagora nello spazio a n dimensioni: il quadrato dell'ipotenusa è uguale alla somma dei quadrati dei cateti. Ne daremo una dimostrazione nel piano che si trasferisce immediatamente al caso n-dimensionale.

Siano f e g due vettori perpendicolari nel piano e consideriamo il triangolo rettangolo costruito su di essi (figura 1). La sua ipotenusa ha la lunghezza del vettore $f + g$. Scriviamo il teorema di Pitagora con la nostra notazione: poiché il quadrato della lunghezza di un vettore è il suo prodotto scalare con sé stesso, il teorema assume la forma

$$(f + g, f + g) = (f, f) + (g, g).$$

La dimostrazione segue subito dalle proprietà del prodotto scalare. Infatti,

$$(f + g, f + g) = (f, f) + (f, g) + (g, f) + (g, g),$$

e i due addendi centrali sono nulli per l'ortogonalità di f e g.

Abbiamo usato soltanto la definizione di lunghezza e di perpendicolarità e le proprietà del prodotto scalare. La dimostrazione non cambia dunque se f e g sono vettori ortogonali dello spazio a n dimensioni. Il teorema di Pitagora è così dimostrato anche in questo spazio.

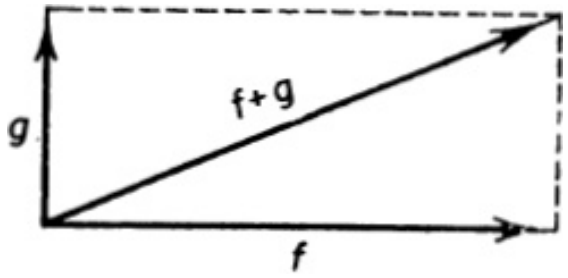


Figura 1.

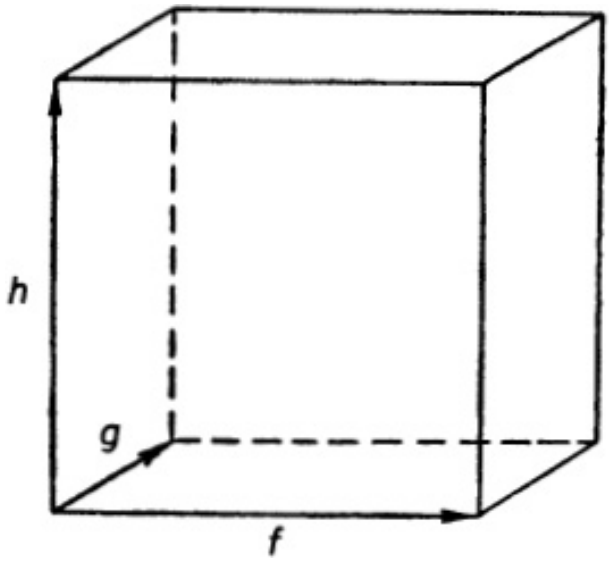


Figura 2.

Se nello spazio a n dimensioni sono dati tre vettori f, g, h a due a due ortogonali, la loro somma $f + g + h$ è la diagonale del parallelepipedo rettangolo costruito su di essi (figura 2) e vale

$$(f + g + h, f + g + h) = (f, f) + (g, g) + (h, h),$$

Ciò significa che il quadrato della lunghezza della diagonale è uguale alla somma dei quadrati delle lunghezze dei tre spigoli uscenti dallo stesso vertice. Lasciamo al lettore la dimostrazione, del tutto analoga alla precedente. Allo stesso modo, per k vettori a due a due ortogonali $f^1, f^2, \ldots, f^k$, vale l'uguaglianza

$$(f^1 + f^2 + \cdots + f^k, f^1 + f^2 + \cdots + f^k) = (f^1, f^1) + (f^2, f^2) + \cdots + (f^k, f^k), \qquad (4)$$

anch'essa facile da dimostrare: il quadrato della diagonale di un «parallelepipedo a k dimensioni» nello spazio a n dimensioni è uguale alla somma dei quadrati degli spigoli che lo generano.

§2. Lo spazio di Hilbert: uno spazio a infinite dimensioni

Collegamento con lo spazio a n dimensioni.

Il concetto di spazio a n dimensioni si è rivelato utile in numerosi problemi matematici e fisici e ha a sua volta stimolato un'ulteriore estensione del concetto di spazio e delle sue applicazioni. I problemi sulle piccole oscillazioni dei sistemi elastici ebbero un ruolo importante nello sviluppo dell'algebra lineare e della geometria degli spazi a n dimensioni.

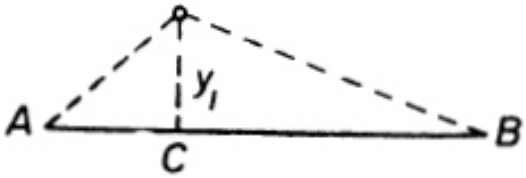


Figura 3.

Consideriamo un esempio classico (figura 3). Sia AB un filo flessibile teso fra A e B, al quale è attaccata una massa in un punto C. Se viene spostata dalla posizione di equilibrio, la massa comincia a oscillare con una frequenza ω, calcolabile conoscendo la tensione del filo, la massa m e la sua posizione. La configurazione del sistema in ogni istante è descritta da un solo numero, lo spostamento y_1 della massa rispetto alla posizione di equilibrio.

Collochiamo ora sul filo AB un numero n di masse nei punti $C_1, C_2, \ldots, C_n$. Supponiamo trascurabile la massa del filo rispetto a quelle applicate. La configurazione è descritta da n numeri $y_1, y_2, \ldots, y_n$, gli spostamenti delle masse dall'equilibrio. È utile sotto molti aspetti considerarli come le componenti di un vettore $(y_1, y_2, \ldots, y_n)$ dello spazio a n dimensioni.

Lo studio delle piccole oscillazioni di questo sistema è strettamente collegato ai fatti fondamentali della geometria n-dimensionale. Per esempio, la determinazione delle frequenze si può ricondurre alla ricerca degli assi di un certo ellissoide nello spazio a n dimensioni.

Consideriamo ora le piccole oscillazioni di una corda tesa fra A e B. Si tratta di una corda idealizzata, cioè di un filo elastico con massa finita distribuita con continuità lungo il filo. La corda si dice omogenea se la densità è costante.

Poiché la massa è distribuita con continuità, non possiamo più descrivere la posizione della corda con un numero finito di dati y_1, ..., y_n: occorre assegnare lo spostamento $y(x)$ di ogni suo punto x. La configurazione in ogni istante è quindi una funzione $y(x)$.

Un filo con n masse applicate alle ascisse x_1, ..., x_n è rappresentato da una spezzata con n vertici interni (figura 4); aumentando le masse, aumenta il numero dei segmenti. Se il numero delle masse cresce indefinitamente, la distanza fra masse adiacenti tende a zero e le masse sono opportunamente distribuite, si ottiene al limite una distribuzione continua: la corda idealizzata. La spezzata si trasforma allora nella curva che descrive la posizione della corda (figura 5).

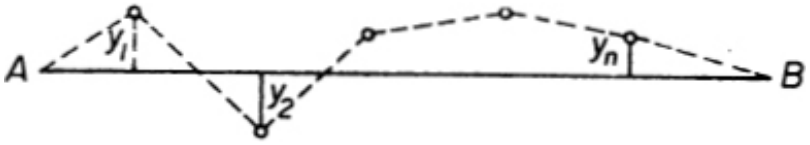


Figura 4.



Figura 5.

Vi è dunque un legame stretto fra le oscillazioni del filo caricato e quelle della corda. Nel primo problema la configurazione è un punto, o vettore, di uno spazio a n dimensioni. È quindi naturale considerare la funzione $f(x)$ che descrive la corda come un vettore, o punto, di uno spazio a infinite dimensioni. Numerosi problemi analoghi conducono alla stessa idea: uno spazio i cui punti siano funzioni $f(x)$ definite su un intervallo.[3]

L'esempio della corda, al quale torneremo nel §4, suggerisce come introdurre i concetti fondamentali di uno spazio a infinite dimensioni.

Lo spazio di Hilbert.

Esamineremo uno dei concetti più diffusi e più importanti per le applicazioni: lo spazio di Hilbert.

Un vettore dello spazio a n dimensioni è una successione di n numeri f_i, con i da 1 a n. Analogamente, nel modello infinito-dimensionale qui considerato un vettore è rappresentato da una funzione $f(x)$, con x variabile fra a e b.

La somma dei vettori e il prodotto per un numero si definiscono come somma delle funzioni e prodotto della funzione per il numero.

La lunghezza del vettore f nello spazio a n dimensioni è definita da

$$\sqrt{\sum_{i=1}^{n} f_i^2}$$

Poiché per le funzioni il ruolo della somma è assunto dall'integrale, la lunghezza del vettore $f(x)$ nello spazio di Hilbert è data da

$$\sqrt{\int_a^b f^2(x)\,dx}. \qquad (5)$$

La distanza fra f e g nello spazio a n dimensioni è la lunghezza di $f - g$, cioè

$$\sqrt{\sum_{i=1}^{n} (f_i - g_i)^2}.$$

Analogamente, la «distanza» fra gli elementi $f(t)$ e $g(t)$ dello spazio funzionale è

$$\sqrt{\int_a^b [f(t) - g(t)]^2\,dt}.$$

L'espressione $\int_a^b [f(t) - g(t)]^2\,dt$ è il quadrato dello scarto quadratico integrale delle funzioni $f(t)$ e $g(t)$. La sua radice quadrata misura la distanza fra i due elementi dello spazio di Hilbert.

Passiamo all'angolo fra vettori. Nello spazio a n dimensioni, l'angolo ϕ fra i vettori non nulli $f = \{f_i\}$ e $g = \{g_i\}$ è definito dalla formula

$$\cos\phi = \frac{\sum_{i=1}^{n} f_i g_i}{\sqrt{\sum_{i=1}^{n} f_i^2}\,\sqrt{\sum_{i=1}^{n} g_i^2}}.$$

Nello spazio a infinite dimensioni sostituiamo le somme con gli integrali corrispondenti e definiamo l'angolo fra due vettori non nulli f e g mediante la formula analoga

$$\cos\phi = \frac{\int_a^b f(t)\,g(t)\,dt}{\sqrt{\int_a^b f^2(t)\,dt}\,\sqrt{\int_a^b g^2(t)\,dt}}. \qquad (6)$$

L'espressione può essere interpretata come coseno di un angolo ϕ purché il valore assoluto della frazione non superi uno, cioè purché

$$\left| \int_a^b f(t)\,g(t)\,dt \right| < \sqrt{\int_a^b f^2(t)\,dt}\,\sqrt{\int_a^b g^2(t)\,dt}. \qquad (7)$$

Questa disuguaglianza vale effettivamente per due funzioni reali a quadrato integrabile $f(t)$ e $g(t)$. Ha un ruolo importante nell'analisi ed è nota come disuguaglianza di Cauchy-Bunjakovskij. Dimostriamola.

Siano $f(x)$ e $g(x)$ due funzioni reali a quadrato integrabile su (a, b), non nulle nello spazio considerato. Scegliamo due numeri reali arbitrari λ e μ e formiamo l'espressione

$$\int_a^b [\lambda f(x) - \mu g(x)]^2 \, dx.$$

Poiché la funzione integranda $[\lambda f(x) - \mu g(x)]^2$ è non negativa, abbiamo

$$\int_a^b [\lambda f(x) - \mu g(x)]^2 \, dx \geqslant 0;$$

ossia

$$2\lambda\mu \int_a^b f(x)\, g(x)\, dx \leqslant \lambda^2 \int_a^b f^2(x)\, dx + \mu^2 \int_a^b g^2(x)\, dx.$$

Per brevità poniamo

$$\left| \int_a^b f(x)\, g(x)\, dx \right| = C, \int_a^b f^2(x)\, dx = A, \int_a^b g^2(x)\, dx = B. \qquad (8)$$

Con questa notazione la disuguaglianza diventa[4]

$$2\lambda\mu C \leqslant \lambda^2 A + \mu^2 B. \qquad (9)$$

Se C = 0, la disuguaglianza da dimostrare è immediata; supponiamo dunque C > 0. Essa vale per qualsiasi λ e μ; possiamo quindi scegliere in particolare

$$\lambda = \sqrt{\frac{C}{A}}, \mu = \sqrt{\frac{C}{B}}. \qquad (10)$$

Sostituendo nella (9), otteniamo

$$\frac{C}{\sqrt{AB}} \leqslant 1.$$

Sostituendo infine ad A, B e C le espressioni (8), ritroviamo la disuguaglianza di Cauchy-Bunjakovskij.

In geometria il prodotto scalare è il prodotto delle lunghezze dei vettori per il coseno dell'angolo compreso. Nel nostro caso le lunghezze di f e g sono

$$\sqrt{(\int_a^b f(t)^2\, dt)} \text{ e } \sqrt{(\int_a^b g(t)^2\, dt)}$$

e il coseno dell'angolo è definito da

$$\cos\phi = \frac{\int_a^b f(x)\, g(x)\, dx}{\sqrt{\int_a^b f^2(x)\, dx}\, \sqrt{\int_a^b g^2(x)\, dx}}.$$

Moltiplicando queste espressioni si ottiene la formula del prodotto scalare di due vettori dello spazio di Hilbert:

$$(f, g) = \int_a^b f(x)\, g(x)\, dx. \qquad (11)$$

Da questa formula risulta che il prodotto scalare di f con sé stesso è il quadrato della sua lunghezza.

Se il prodotto scalare di due vettori non nulli f e g è zero, allora $\cos \phi = 0$ e l'angolo definito fra essi è 90°. Perciò le funzioni che soddisfano

$$(f, g) = \int_a^b f(x)\, g(x)\, dx = 0,$$

si dicono ortogonali.

Il teorema di Pitagora del §1 vale anche nello spazio di Hilbert. Siano $f_1(x)$, $f_2(x)$, ..., $f_N(x)$ funzioni a due a due ortogonali e sia

$$f(x) = f_1(x) + f_2(x) + \cdots + f_N(x).$$

Il quadrato della lunghezza di f è allora la somma dei quadrati delle lunghezze di f_1, ..., f_N.

Poiché le lunghezze nello spazio di Hilbert sono espresse mediante integrali, il teorema assume la forma

$$\int_a^b f^2(x)\, dx = \int_a^b f_1^2(x)\, dx + \int_a^b f_2^2(x)\, dx + \cdots + \int_a^b f_N^2(x)\, dx. \qquad (12)$$

La dimostrazione non differisce da quella data nel §1 per lo spazio a n dimensioni.

Non abbiamo ancora precisato quali funzioni siano ammesse come vettori dello spazio di Hilbert. Dobbiamo prendere quelle per cui $\int_a^b f^2(x)\, dx$ esiste ed è finito. Potrebbe sembrare naturale limitarsi alle funzioni continue sull'intervallo chiuso, per le quali $\int_a^b f^2(x)\, dx$ esiste sempre. La teoria diventa però più completa e naturale interpretando $\int_a^b f^2(x)\, dx$ nel senso generalizzato dell'integrale di Lebesgue (capitolo XV). Si ottiene così lo spazio delle funzioni misurabili a quadrato integrabile, identificando le funzioni uguali quasi ovunque.

Questa estensione dell'integrale, e quindi della classe delle funzioni considerate, è necessaria all'analisi funzionale come una teoria rigorosa dei numeri reali è necessaria ai fondamenti del calcolo differenziale e integrale. La generalizzazione dell'integrale, nata all'inizio del Novecento nello sviluppo della teoria delle funzioni di variabile reale, si rivelò dunque essenziale per l'analisi funzionale e per i settori a essa collegati.

§3. Sviluppi in sistemi ortogonali di funzioni

Definizione ed esempi di sistemi ortogonali.

Scelti nel piano due vettori unitari perpendicolari e_1 ed e_2 (figura 7), ogni vettore del piano si può scomporre lungo le loro direzioni, cioè scrivere nella forma

$$f = a_1 e_1 + a_2 e_2 ,$$

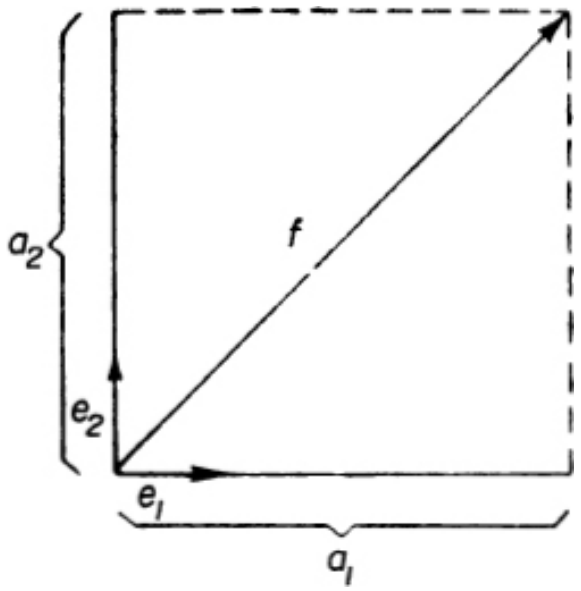


Figura 7.

dove a_1 e a_2 sono le proiezioni, con segno, di f sugli assi orientati da e_1 ed e_2. Poiché la proiezione è il prodotto della lunghezza del vettore per il coseno dell'angolo con l'asse, dalla definizione di prodotto scalare segue

$$a_1 = (f, e_1),$$
$$a_2 = (f, e_2).$$

Analogamente, scelti nello spazio tridimensionale tre vettori unitari a due a due perpendicolari e_1, e_2, e_3, ogni vettore f si scrive

$$f = a_1 e_1 + a_2 e_2 + a_3 e_3 ,$$

dove

$$a_k = (f, e_k) \ (k = 1, 2, 3).$$

Anche nello spazio di Hilbert possiamo considerare sistemi di vettori a due a due ortogonali, ossia funzioni $\phi_1(x)$, $\phi_2(x)$, ..., $\phi_n(x)$, ...

Questi sistemi, detti ortogonali, hanno un ruolo importante nell'analisi. Compaiono in problemi molto diversi di fisica matematica, equazioni integrali, calcolo approssimato, teoria delle funzioni di variabile reale e così via. L'ordinamento e l'unificazione dei concetti relativi a

questi sistemi furono fra le ragioni che condussero, all'inizio del Novecento, al concetto generale di spazio di Hilbert.

Diamo una definizione precisa. Un sistema di funzioni

$$\phi_1(x), \phi_2(x), \cdots, \phi_n(x), \cdots$$

si dice ortogonale se due funzioni distinte qualsiasi del sistema sono ortogonali, cioè se

$$\int_a^b \phi_i(x)\phi_k(x)\,dx = 0 \text{ per } i \neq k. \quad (13)$$

Nello spazio tridimensionale avevamo richiesto che i vettori fossero unitari. Ricordando la definizione di lunghezza, nello spazio di Hilbert questa condizione si scrive

$$\int_a^b \phi_k^2(x)\,dx = 1. \quad (14)$$

Un sistema che soddisfa le condizioni (13) e (14) si dice ortonormale.

Vediamo alcuni esempi.

1. Nell'intervallo $(-\pi, \pi)$ consideriamo la successione

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots .$$

Due funzioni distinte della successione sono ortogonali, come si verifica calcolando gli integrali corrispondenti. Il quadrato della lunghezza di un vettore è l'integrale del quadrato della funzione; per i vettori della successione

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots$$

tali quadrati sono quindi gli integrali

$$\int_{-\pi}^{\pi} dx = 2\pi, \quad \int_{-\pi}^{\pi} \cos^2 nx\,dx = \pi, \quad \int_{-\pi}^{\pi} \sin^2 nx\,dx = \pi,$$

I vettori sono dunque ortogonali, ma non normalizzati. Il primo ha lunghezza $\sqrt{2\pi}$, tutti gli altri lunghezza $\sqrt{\pi}$. Dividendo ciascun vettore per la propria lunghezza, otteniamo il sistema ortonormale trigonometrico

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \cdots.$$

Storicamente questo è uno dei primi e più importanti esempi di sistemi ortogonali. Comparve nei lavori di Eulero, D. Bernoulli e d'Alembert sui problemi delle corde vibranti; il suo studio ebbe un ruolo essenziale nello sviluppo dell'intera analisi.[5]

Il suo legame con le corde vibranti non è casuale. I problemi sulle piccole oscillazioni dei mezzi elastici conservativi conducono a sistemi di funzioni ortogonali che descrivono le oscillazioni proprie del sistema (§4). Per esempio, nello studio delle oscillazioni di una sfera compaiono le funzioni sferiche; per una membrana circolare o un cilindro compaiono le funzioni cilindriche, e così via.

2. Possiamo costruire un sistema ortogonale le cui funzioni siano tutte polinomi: la successione dei polinomi di Legendre

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n(x^2 - 1)^n}{dx^n},$$

Dunque $P_n(x)$ è, a meno di un fattore costante, la derivata n-esima di $(x^2 - 1)^n$. Scriviamo i primi polinomi:

$$P_0(x) = 1;$$
$$P_1(x) = x;$$
$$P_2(x) = \tfrac{1}{2}(3x^2 - 1);$$
$$P_3(x) = \tfrac{1}{2}(5x^3 - 3x).$$

Evidentemente $P_n(x)$ ha grado n. Lasciamo al lettore la verifica che questi polinomi sono ortogonali nell'intervallo $(-1, 1)$.

La teoria generale dei polinomi ortogonali, anche rispetto a una funzione peso, fu sviluppata nella seconda metà dell'Ottocento dal celebre matematico russo P. L. Čebyšëv.

Sviluppi in sistemi ortogonali.

Come nello spazio tridimensionale ogni vettore si esprime come combinazione lineare di tre vettori unitari a due a due ortogonali,

$$f = a_1 e_1 + a_2 e_2 + a_3 e_3,$$

così nello spazio funzionale si pone il problema di sviluppare una funzione arbitraria f in serie rispetto a un sistema ortonormale, cioè di rappresentarla nella forma

$$f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x) + \cdots. \qquad (15)$$

La convergenza della serie (15) a f va intesa rispetto alla distanza nello spazio di Hilbert: lo scarto quadratico integrale fra la somma parziale

$$S_n(t) = \sum_{k=1}^{n} a_k\phi_k(t)$$

e la funzione $f(t)$ deve tendere a zero per $n \to \infty$, cioè

$$\lim_{n\to\infty} \int_a^b [f(t) - S_n(t)]^2\, dt = 0. \qquad (16)$$

Questa convergenza si chiama convergenza in media quadratica.

Gli sviluppi in sistemi ortogonali ricorrono frequentemente nell'analisi e costituiscono un metodo importante per risolvere problemi di fisica matematica. Se, per esempio, si considera il sistema trigonometrico nell'intervallo $(-\pi, \pi)$,

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots,$$

si ottiene il classico sviluppo in serie trigonometrica[6]

$$f(x) = a_0/2 + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

Supponiamo che lo sviluppo (15) sia possibile per ogni funzione *f* dello spazio di Hilbert e determiniamone i coefficienti a_n. Moltiplichiamo scalarmente entrambi i membri per una funzione ϕ_m del sistema. Otteniamo

$$(f, \phi_m) = a_1 (\phi_1 , \phi_m) + a_2(\phi_2 , \phi_m) + \cdots + a_m(\phi_m , \phi_m) + a_{m+1}(\phi_{m+1} , \phi_m) + \cdots,$$

Poiché $(\phi_m, \phi_n) = 0$ per $m \neq n$ e $(\phi_m, \phi_m) = 1$, il coefficiente a_m è determinato da

$$a_m = (f, \phi_m) \ (m = 1, 2, \cdots).$$

Come nello spazio tridimensionale, i coefficienti a_m sono dunque le proiezioni del vettore *f* lungo i vettori ϕ_m.

Dalla definizione di prodotto scalare segue che i coefficienti dello sviluppo di $f(x)$ nel sistema ortonormale $\phi_1(x)$, $\phi_2(x)$, ...,

$$f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x) + \cdots \qquad (17)$$

sono dati dalle formule

$$a_m = \int_a^b f(t)\, \phi_m(t)\, dt. \qquad (18)$$

Consideriamo, per esempio, il sistema trigonometrico ortonormale già introdotto:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \cdots.$$

Abbiamo allora

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

dove

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx\, dx,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx\, dx.$$

Abbiamo così ottenuto le formule per calcolare i coefficienti dello sviluppo trigonometrico, supponendo naturalmente che tale sviluppo sia possibile.[7]

Abbiamo determinato i coefficienti dello sviluppo (18) nell'ipotesi che esso esista. Ma un sistema ortogonale infinito ϕ_1, ϕ_2, ... può non bastare a sviluppare ogni funzione dello spazio di Hilbert. Occorre una condizione ulteriore, detta completezza del sistema.

Un sistema ortogonale si dice completo se non è possibile aggiungervi neppure una funzione non nulla nello spazio di Hilbert che sia ortogonale a tutte quelle del sistema.

È facile costruire un sistema non completo. Prendiamo un sistema ortogonale, per esempio quello trigonometrico, ed eliminiamo una funzione, come cos *x*. Il sistema infinito rimanente

$$1, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots$$

è ancora ortogonale, ma non completo, perché cos *x* è ortogonale a tutte le sue funzioni.

Se un sistema non è completo, non ogni funzione può essere sviluppata in esso. Tentando infatti di sviluppare una funzione non nulla $f_0(x)$ ortogonale a tutte le funzioni del sistema, la (18) darebbe coefficienti tutti nulli, mentre f_0 non è il vettore nullo.

Vale il seguente teorema: dato un sistema ortonormale completo $\phi_1(x)$, $\phi_2(x)$, ... nello spazio di Hilbert, ogni funzione $f(x)$ dello spazio si sviluppa in serie, convergente in media quadratica, nelle funzioni del sistema:[8]

$$f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x) + \cdots.$$

I coefficienti a_n sono le proiezioni di *f* sugli elementi del sistema ortonormale:

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

Il teorema di Pitagora nello spazio di Hilbert, stabilito nel §2, permette di trovare una relazione interessante fra i coefficienti a_k e $f(x)$. Indichiamo con $r_n(x)$ la differenza fra la funzione e la somma dei primi *n* termini della serie:

$$r_n(x) = f(x) - [a_1\phi_1(x) + \cdots + a_n\phi_n(x)].$$

La funzione $r_n(x)$ è ortogonale a $\phi_1(x)$, ..., $\phi_n(x)$. Verifichiamolo, per esempio, per ϕ_1, mostrando che $\int_a^b r_n(x)\,\phi_1(x)\,dx = 0$. Abbiamo

$$\begin{aligned}&\int_a^b r_n(x)\phi_1(x)\,dx\\ &= \int_a^b [f(x) - a_1\phi_1(x) - \ldots - a_n\phi_n(x)]\phi_1(x)\,dx\\ &= \int_a^b f(x)\phi_1(x)\,dx - a_1\int_a^b \phi_1(x)^2\,dx.\end{aligned}$$

[9]

Poiché $a_1 = \int_a^b f(x)\,\phi_1(x)\,dx$ e $\int_a^b \phi_1^2(x)\,dx = 1$, segue che $\int_a^b r_n(x)\,\phi_1(x)\,dx = 0$.

Nell'uguaglianza

$$f(x) = a_1\phi_1(x) + a_2\phi_2(x) + \cdots + a_n\phi_n(x) + r_n(x) \qquad (19)$$

gli addendi del secondo membro sono dunque a due a due ortogonali. Per il teorema di Pitagora, il quadrato della lunghezza di *f* è la somma dei quadrati delle loro lunghezze:

$$\int_a^b f^2(x)\,dx = \int_a^b [a_1\phi_1(x)]^2\,dx + \cdots + \int_a^b [a_n\phi_n(x)]^2\,dx + \int_a^b r_n^2(x)\,dx.$$

Poiché il sistema è normalizzato, come nella (14), abbiamo

$$\int_a^b f^2(x)\,dx = a_1^2 + a_2^2 + \cdots + a_n^2 + \int_a^b r_n^2(x)\,dx. \qquad (20)$$

La serie $\sum_{k=1}^{\infty} a_k\phi_k(x)$ converge in media quadratica. Ciò significa che

$$\int_a^b [f(x) - a_1\phi_1(x) - \cdots - a_n\phi_n(x)]^2\,dx \to 0,$$

ossia

[10]

$$\int_a^b r_n^2(x)\,dx \to 0.$$

Dalla formula (20) otteniamo allora

$$\sum_{k=1}^{\infty} a_k^2 = \int_a^b f^2(x)\,dx, \qquad (21)$$

L'integrale del quadrato della funzione è dunque la somma dei quadrati dei coefficienti del suo sviluppo nel sistema ortonormale completo. Questa identità è detta uguaglianza di Parseval; la sua validità per ogni funzione dello spazio di Hilbert esprime una condizione equivalente alla completezza del sistema.

Richiamiamo l'attenzione su un problema importante: quali numeri a_k possono essere coefficienti dello sviluppo di una funzione dello spazio di Hilbert? La (21) impone che la serie $\sum_{k=1}^{\infty} a_k^2$ converga. Questa condizione è anche sufficiente: una successione a_k è la successione dei coefficienti di una funzione rispetto a un sistema ortonormale se e solo se converge $\sum_{k=1}^{\infty} a_k^2$.

Questo teorema fondamentale vale quando lo spazio di Hilbert comprende tutte le funzioni a quadrato integrabile secondo Lebesgue, con l'identificazione quasi ovunque (§2). Limitandosi, per esempio, alle funzioni continue, il problema di caratterizzare le successioni dei coefficienti diventerebbe inutilmente complicato.

È questa soltanto una delle ragioni che hanno condotto a usare l'integrale di Lebesgue nella definizione dello spazio di Hilbert.

§4. Le equazioni integrali

In questa sezione presentiamo uno dei settori più importanti e storicamente più antichi dell'analisi funzionale: la teoria delle equazioni integrali, che ebbe un ruolo essenziale anche nei suoi sviluppi successivi. Oltre alle esigenze interne della matematica, come i problemi al contorno per le equazioni alle derivate parziali del capitolo VI, furono determinanti numerosi problemi fisici. Nel Novecento le equazioni integrali, accanto a quelle differenziali, sono diventate uno degli strumenti principali per lo studio matematico della fisica. Esporremo alcuni risultati della teoria, strettamente collegati e nati, direttamente o indirettamente, dallo studio delle piccole oscillazioni dei sistemi elastici.

Il problema delle piccole oscillazioni dei sistemi elastici.

Torniamo al problema del §2 e cerchiamo le equazioni che descrivono le oscillazioni. Per semplicità consideriamo un sistema elastico lineare, come una corda di lunghezza l (figura 8) o un'asta elastica (figura 9). Supponiamo che all'equilibrio il sistema si trovi sul segmento Ol dell'asse x. Applichiamo una forza unitaria nel punto x: tutti i punti subiscono un certo spostamento. Indichiamo con $k(x, y)$ quello prodotto nel punto y (figura 8).

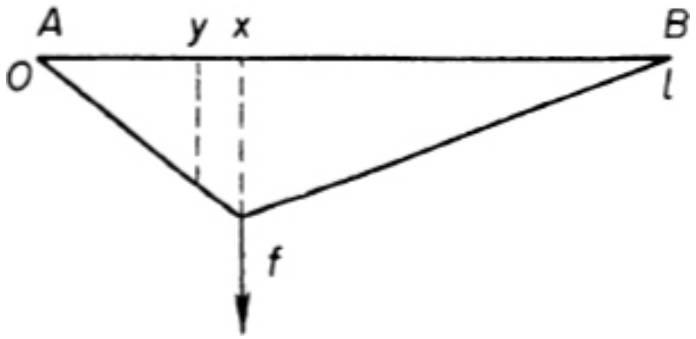


Figura 8.

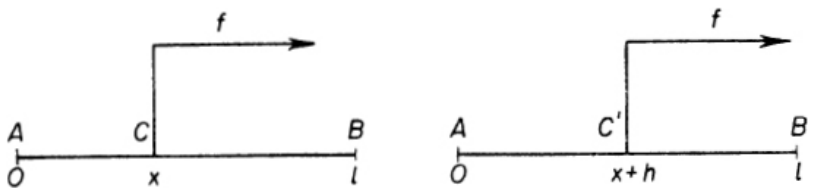


Figura 9.

La funzione $k(x, y)$ dipende da due punti: quello x in cui si applica la forza e quello y in cui si misura lo spostamento. Si chiama funzione d'influenza, o funzione di Green.

Per il sistema elastico lineare conservativo, dal principio di conservazione dell'energia si deduce una proprietà importante della funzione di Green, la legge di reciprocità: lo spostamento prodotto in y da una forza applicata in x è uguale allo spostamento prodotto in x dalla stessa forza applicata in y. In altre parole,

$$k(x, y) = k(y, x). \qquad (22)$$

Calcoliamo, per esempio, la funzione di Green per gli spostamenti longitudinali di un'asta elastica; nella figura 8 erano rappresentati invece spostamenti trasversali. Consideriamo un'asta AB di lunghezza l, fissata agli estremi (figura 9). Applichiamo in C una forza f diretta verso B. L'asta si deforma e C si sposta in C'; indichiamo con h l'entità dello spostamento. Determiniamo h, dal quale ricaveremo poi lo spostamento di un punto arbitrario y. Useremo la legge di Hooke: la forza è proporzionale all'allungamento relativo, cioè al rapporto fra allungamento e lunghezza iniziale. Una relazione analoga vale per la compressione.

La forza f allunga il tratto AC, generando una reazione di intensità T_1; contemporaneamente comprime CB, generando una reazione T_2. Per la legge di Hooke,

$$T_1 = \kappa \frac{h}{x}, \; T_2 = \kappa \frac{h}{l - x},$$

dove κ è il coefficiente di proporzionalità che caratterizza la rigidezza dell'asta. L'equilibrio delle forze in C dà

$$f = \kappa h/x + \kappa h/(l - x), \text{ ossia } f = \kappa l h/[x(l - x)].$$

Da qui

$$h = \frac{f}{\kappa l} x(l - x).$$

Per trovare lo spostamento di un punto y del tratto AC, con $y < x$, osserviamo che, per la legge di Hooke, l'allungamento relativo è uniforme: il rapporto fra lo spostamento di un punto e la sua distanza dall'estremo fisso non dipende dal punto scelto. Indichiamo lo spostamento di y con k.

Confrontando gli spostamenti relativi in x e in y, otteniamo

$$\frac{k}{y} = \frac{h}{x};$$

quindi

$$k = hy/x = fy(l - x)/(\kappa l), \text{ per } y < x.$$

Analogamente, per un punto del tratto CB, con $y > x$, si ottiene

$$k = h \frac{l - y}{l - x} = \frac{f}{\kappa l} x(l - y).$$

Ricordando che $k(x, y)$ è lo spostamento in y dovuto a una forza unitaria applicata in x, la funzione di Green dell'asta ha dunque la forma

$$k(x, y) = y(l - x)/(\kappa l), \text{ per } y \leq x;$$
$$k(x, y) = x(l - y)/(\kappa l), \text{ per } y \geq x.$$

In modo analogo si trova la funzione di Green della corda. Se la tensione è T e la lunghezza l, una forza unitaria applicata in x produce, nell'approssimazione dei piccoli spostamenti, la forma della figura 8; lo spostamento in y è

$$k(x, y) = x(l - y)/(Tl), \text{ per } x \leq y;$$
$$k(x, y) = y(l - x)/(Tl), \text{ per } x \geq y.$$

La forma è la stessa di quella ricavata per l'asta, con T al posto di κ.

Con la funzione di Green possiamo esprimere lo spostamento prodotto da una forza distribuita con densità $f(y)$. Su un intervallo di lunghezza Δy agisce la forza $f(y)\Delta y$, che possiamo considerare approssimativamente concentrata in y. Per reciprocità, essa produce in x lo spostamento $k(x, y)f(y)\Delta y$. Lo spostamento dovuto all'intero carico è approssimativamente

$$\sum k(x, y)\, f(y)\, \Delta y.$$

Passando al limite per $\Delta y \to 0$, lo spostamento $u(x)$ dovuto al carico distribuito è

$$u(x) = \int_a^b k(x, y)\, f(y)\, dy. \qquad (23)$$

Supponiamo ora che non agiscano forze esterne. Se il sistema viene spostato dall'equilibrio e lasciato libero, comincia a muoversi: questi moti si chiamano oscillazioni libere.

Scriviamone l'equazione mediante $k(x, y)$. Indichiamo con $u(x, t)$ lo spostamento del punto x all'istante t. L'accelerazione è allora $\partial^2 u(x, t)/\partial t^2$.

Sia ρ la densità lineare uniforme, così che $\rho\, dy$ è la massa di un elemento di lunghezza dy. Per le leggi della meccanica otteniamo l'equazione del moto sostituendo nella (23) alla forza $f(y)dy$ il prodotto della massa per l'accelerazione, preso con segno opposto: $-[\partial^2 u(y, t)/\partial t^2]\rho\, dy$.

L'equazione delle oscillazioni libere è pertanto

$$u(x, t) = -\int_a^b k(x, y)\, \frac{\partial^2 u(y, t)}{\partial t^2}\, \rho\, dy.$$

Hanno particolare importanza le oscillazioni armoniche, cioè i moti per cui

$$u(x, t) = u(x) \sin \omega t.$$

Ogni punto compie un'oscillazione sinusoidale con una certa pulsazione ω, uguale per tutti i punti x. Nel seguito, seguendo l'uso del testo, la chiameremo anche frequenza angolare.

Vedremo che ogni oscillazione libera del sistema considerato si compone di oscillazioni armoniche.

Poniamo

$$u(x, t) = u(x) \sin \omega t$$

nell'equazione del moto e semplifichiamo il fattore sin ωt. Per determinare $u(x)$ otteniamo

$$u(x) = \rho\omega^2 \int_a^b k(x, y)\, u(y)\, dy. \qquad (24)$$

Questa è un'equazione integrale omogenea per la funzione $u(x)$.

Per ogni ω la (24) ha la soluzione ovvia $u(x) \equiv 0$, che corrisponde alla quiete. I valori di ω per cui esistono soluzioni non nulle si chiamano frequenze proprie del sistema.

Poiché le soluzioni non nulle non esistono per ogni ω, le oscillazioni libere armoniche sono possibili soltanto a frequenze determinate. La più piccola è la frequenza fondamentale; le altre sono le frequenze dei toni superiori.

Per i sistemi continui qui considerati esiste una successione infinita di frequenze proprie, detta spettro delle frequenze:

$$\omega_1 , \omega_2 , \cdots, \omega_n , \cdots$$

La soluzione non nulla $u_n(x)$ della (24), associata alla frequenza ω_n, determina la forma del corrispondente modo proprio di oscillazione.

Per esempio, per una corda uniforme tesa e fissata nei punti O e l, le frequenze angolari proprie sono

$$a\frac{\pi}{l} , 2a\frac{\pi}{l} , 3a\frac{\pi}{l} , \cdots, na\frac{\pi}{l} , \cdots,$$

dove a dipende dalla densità e dalla tensione, precisamente $a = \sqrt{T/\rho}$. La frequenza fondamentale è $\omega_1 = a\pi/l$; le altre sono $\omega_2 = 2\omega_1$, $\omega_3 = 3\omega_1$, ..., $\omega_n = n\omega_1$. Le forme dei modi corrispondenti sono date da

$$u_n(x) = \sin \frac{n\pi}{l} x$$

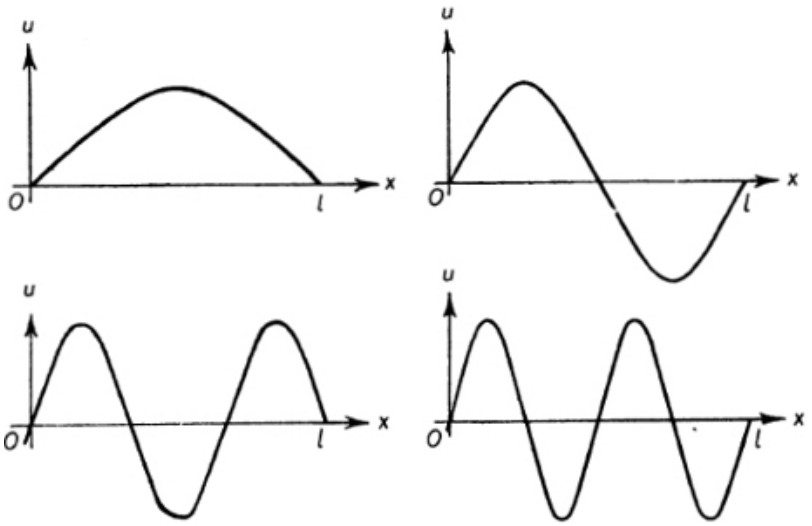


Figura 10.

e sono rappresentate per $n = 1, 2, 3, 4$ nella figura 10.

Finora abbiamo considerato oscillazioni libere. Se invece agisce una forza esterna armonica, la determinazione della risposta armonica conduce alla seguente equazione integrale non omogenea per $u(x)$:

$$u(x) = \rho\omega^2 \int_a^b k(x, y)\, u(y)\, dy + h(x). \qquad (25)$$

Proprietà delle equazioni integrali.

Abbiamo incontrato esempi delle equazioni integrali

$$f(x) = \lambda \int_a^b k(x, y)\, f(y)\, dy \qquad (26)$$

e

$$f(x) = \lambda \int_a^b k(x, y)\, f(y)\, dy + h(x), \qquad (27)$$

La prima nasce dalle oscillazioni libere di un sistema elastico; la seconda dalle oscillazioni forzate, cioè prodotte da forze esterne.

L'incognita è $f(x)$. La funzione assegnata $k(x, y)$ si chiama nucleo dell'equazione. La (27) è un'equazione integrale lineare non omogenea; la (26) è la corrispondente equazione omogenea, ottenuta ponendo $h(x) = 0$.

L'equazione omogenea ha sempre la soluzione nulla. Fra le soluzioni dell'equazione omogenea e di quella non omogenea esiste un legame stretto. Citiamo, per esempio, il seguente teorema, nelle usuali ipotesi di compattezza dell'operatore integrale: se l'equazione omogenea ha soltanto la soluzione nulla, quella non omogenea ammette una soluzione per ogni funzione $h(x)$ dello spazio considerato.

Se, per un certo valore λ, l'equazione omogenea ha una soluzione non nulla $f(x)$, il parametro λ si chiama qui valore caratteristico dell'equazione, e f autofunzione. Per le oscillazioni libere esso è legato alla frequenza dalla relazione $\lambda = \rho\omega^2$; le autofunzioni descrivono le forme dei modi armonici. Il valore caratteristico di questa scrittura è il reciproco dell'autovalore del corrispondente operatore integrale.

Nei problemi sulle oscillazioni, la conservazione dell'energia conduce alla relazione

$$k(x, y) = k(y, x). \qquad (28)$$

Un nucleo che soddisfa la (28) si dice simmetrico.

Per un nucleo reale simmetrico che definisca un operatore compatto, autofunzioni e valori caratteristici hanno importanti proprietà. I valori caratteristici sono reali e, se l'operatore ha rango infinito, formano una successione infinita:

$$\lambda_1 , \lambda_2 , \cdots, \lambda_n , \cdots .$$

A ogni valore caratteristico corrisponde uno spazio di autofunzioni di dimensione finita, eventualmente superiore a uno. Autofunzioni associate a valori distinti sono sempre ortogonali.[11]

Scegliendo una base ortonormale in ciascuno di questi spazi, otteniamo quindi un sistema ortogonale di autofunzioni. Quando è completo, ossia quando ogni funzione dello spazio di Hilbert può essere sviluppata in esso? In particolare, se l'equazione

$$\int_a^b k(x, y) f(y)\, dy = 0 \qquad (29)$$

è soddisfatta soltanto da $f(y) = 0$, il sistema delle autofunzioni dell'equazione integrale

$$\lambda \int_a^b k(x, y) f(y)\, dy = f(x)$$

è completo.[12]

In questo caso ogni funzione a quadrato integrabile si sviluppa in serie di autofunzioni. Lo studio delle equazioni integrali offre così un metodo generale e potente per dimostrare la completezza di importanti sistemi ortogonali e l'esistenza degli sviluppi corrispondenti. Si può dimostrare in questo modo la completezza dei sistemi trigonometrici, delle funzioni cilindriche, delle funzioni sferiche e di molti altri sistemi.

Nei problemi sulle oscillazioni, la possibilità di sviluppare una funzione in autofunzioni significa che ogni oscillazione può essere decomposta in modi armonici, con le rispettive dipendenze temporali. Ne deriva un metodo largamente applicabile in meccanica e fisica: oscillazioni dei corpi elastici, acustica, onde elettromagnetiche e così via.

La teoria delle equazioni integrali lineari diede impulso alla teoria generale degli operatori lineari, della quale costituisce una parte organica. Nei decenni cui si riferisce l'opera, i metodi

generali degli operatori contribuirono a loro volta vigorosamente allo sviluppo delle equazioni integrali.

§5. Operatori lineari e ulteriori sviluppi dell'analisi funzionale

Nella sezione precedente abbiamo visto che le oscillazioni di un sistema elastico conducono alla ricerca di valori caratteristici e autofunzioni di equazioni integrali. Gli stessi problemi si possono ricondurre allo studio degli autovalori e delle autofunzioni di equazioni differenziali lineari.[13] Molti altri problemi fisici richiedono calcoli analoghi.

Consideriamo un altro esempio. Nella radiotecnica si usano ampiamente guide d'onda per trasmettere oscillazioni elettromagnetiche ad alta frequenza: sono tubi metallici cavi nei quali si propagano le onde. In una guida d'onda cava possono propagarsi soltanto onde con lunghezza d'onda non superiore a un certo limite. La determinazione della lunghezza d'onda critica si riconduce agli autovalori di un'equazione differenziale.

Problemi agli autovalori compaiono inoltre nell'algebra lineare, nelle equazioni differenziali ordinarie, nello studio della stabilità e in molti altri settori.

Divenne quindi necessario esaminare questi problemi collegati da un punto di vista unitario: quello della teoria generale degli operatori lineari. Molti problemi concreti di autovalori e autofunzioni furono compresi pienamente soltanto alla luce di questa teoria, che si rivelò perciò uno strumento di ricerca assai fecondo nei settori ai quali si applica.

La meccanica quantistica, che usa ampiamente i metodi operatoriali, ebbe un ruolo decisivo nei successivi sviluppi. Il suo apparato matematico fondamentale è la teoria degli operatori autoaggiunti. La formulazione dei problemi matematici della meccanica quantistica costituì, e continua a costituire nella prospettiva del testo, un forte stimolo per l'analisi funzionale.

Il punto di vista operatoriale sulle equazioni differenziali e integrali si è rivelato estremamente utile anche per sviluppare metodi pratici di soluzione approssimata.

Concetti fondamentali della teoria degli operatori.

Passiamo ora alle definizioni e ai fatti essenziali.

Nell'analisi abbiamo incontrato il concetto di funzione: nella forma più semplice, una corrispondenza che associa a ogni numero x, valore della variabile indipendente, un numero y, valore della funzione. Lo sviluppo dell'analisi ha richiesto corrispondenze più generali.

Ne troviamo un esempio nel calcolo delle variazioni (capitolo VIII), dove a una funzione si associa un numero. Una tale corrispondenza si chiama funzionale. Per esempio, possiamo associare a una funzione sufficientemente regolare $y = f(x)\ (a \leqslant x \leqslant b)$ la lunghezza dell'arco del suo grafico. Un altro funzionale associa a ogni funzione integrabile $y = f(x)\ (a \leqslant x \leqslant b)$ il suo integrale definito $\int_a^b f(x)\,dx$.

Se consideriamo $f(x)$ come punto di uno spazio a infinite dimensioni, un funzionale è semplicemente una funzione dei punti di questo spazio. Da tale punto di vista, il calcolo delle variazioni ricerca massimi e minimi di funzioni definite su spazi a infinite dimensioni.

Per definire un funzionale continuo dobbiamo prima precisare cosa significhi che due punti dello spazio sono vicini. Nel §2 abbiamo definito la distanza fra due funzioni mediante

$$\sqrt{\int_a^b [f(x) - g(x)]^2 \, dx}.$$

Questa distanza è molto usata, ma non è l'unica possibile. Altri problemi richiedono altri modi di misurare la vicinanza. Per esempio, nella teoria dell'approssimazione delle funzioni (capitolo XII, §3) si usa anche la distanza

$$\max |f(x) - g(x)|.$$

Nel calcolo delle variazioni si usano ancora altri modi di misurare la distanza fra funzioni. Definizioni diverse della distanza conducono a spazi a infinite dimensioni diversi.

Gli spazi funzionali differiscono dunque sia per l'insieme delle funzioni sia per la distanza assegnata. Se, per esempio, prendiamo tutte le funzioni a quadrato integrabile, identificate quasi ovunque, e poniamo

$$\sqrt{\int_a^b [f(x) - g(x)]^2 \, dx},$$

otteniamo lo spazio di Hilbert del §2. Se invece prendiamo tutte le funzioni continue su un intervallo chiuso e poniamo la distanza uguale a $\max |f(x) - g(x)|$, otteniamo lo spazio C.

Nello studio delle equazioni integrali incontriamo espressioni della forma

$$g(x) = \int_a^b k(x, y) f(y) \, dy.$$

Fissato il nucleo $k(x, y)$, questa formula stabilisce una regola che associa a una funzione f un'altra funzione g.

Una corrispondenza di questo tipo si chiama operatore.

Un operatore A nello spazio di Hilbert è una regola che associa a una funzione f una funzione g dello stesso spazio. La corrispondenza può non essere definita per tutte le funzioni: l'insieme di quelle per cui esiste $g = Af$ si chiama dominio di definizione dell'operatore, analogamente al dominio di una funzione nell'analisi ordinaria. Scriviamo

$$g = Af. \qquad (30)$$

L'operatore si dice lineare se il suo dominio è un sottospazio e alla somma di f_1 e f_2 associa la somma di Af_1 e Af_2, mentre al prodotto λf associa λAf, cioè

$$A(f_1 + f_2) = Af_1 + Af_2 \qquad (31)$$

e

$$A(\lambda f) = \lambda A f. \qquad (32)$$

Talvolta si richiede anche la continuità: se f_n converge a f nel dominio, allora Af_n deve convergere ad Af, rispetto alle distanze degli spazi considerati.

Vediamo alcuni esempi di operatori lineari.

1. Associamo a ogni funzione $f(x)$ la funzione $g(x) = \int_a^x f(t)\,dt$, cioè una sua primitiva con valore iniziale fissato a zero. La linearità segue dalle proprietà usuali dell'integrale: l'integrale di una somma è la somma degli integrali e un fattore costante può essere portato fuori dal segno d'integrale.

2. Associamo a ogni funzione differenziabile $f(x)$ la derivata $f'(x)$. Questo operatore si indica solitamente con D:

$$f'(x) = D\,f(x).$$

L'operatore non è definito su tutto lo spazio di Hilbert, ma su un opportuno dominio di funzioni la cui derivata appartiene ancora allo spazio. Nella formulazione usuale in L^2 si usano funzioni assolutamente continue con derivata a quadrato integrabile, eventualmente soggette a condizioni al contorno.

3. I primi due esempi riguardano spazi a infinite dimensioni, ma abbiamo incontrato operatori lineari anche in dimensione finita. Nel capitolo III abbiamo studiato le trasformazioni affini: una trasformazione affine del piano o dello spazio che lasci fissa l'origine è un operatore lineare in dimensione due o tre. Le trasformazioni lineari del capitolo XVI sono operatori lineari nello spazio a n dimensioni.

4. Le equazioni integrali ci hanno già fornito una classe di operatori lineari molto importante e ricca di applicazioni: gli operatori integrali. Fissiamo una funzione $k(x, y)$; la formula

$$g(x) = \int_a^b k(x, y) f(y)\,dy$$

associa a ogni funzione ammissibile f una funzione g. Simbolicamente,

$$g = Af.$$

L'operatore A si chiama integrale. Potremmo citare molti altri esempi importanti di operatori integrali.

Nel §4 abbiamo considerato l'equazione non omogenea

$$f(x) = \lambda \int_a^b k(x, y) f(y)\,dy + h(x).$$

Nel linguaggio degli operatori essa si scrive

$$f = \lambda A f + h, \qquad (33)$$

dove λ è un numero assegnato, h una funzione assegnata, cioè un vettore dello spazio a infinite dimensioni, e f l'incognita. L'equazione omogenea diventa

$$f = \lambda A f. \qquad (34)$$

I teoremi classici sulle equazioni integrali, come quello del §4 che collega la risolubilità dell'equazione non omogenea a quella dell'omogenea, non valgono per qualsiasi equazione operatoriale. Si possono però indicare condizioni generali su A che ne garantiscano la validità.

Queste condizioni si esprimono in termini topologici: A deve trasformare la palla unitaria, cioè l'insieme dei vettori di lunghezza non superiore a 1, in un insieme relativamente compatto, ossia con chiusura compatta. Si parla allora di operatore compatto.

Autovalori e autovettori degli operatori.

Il problema dei valori caratteristici e delle autofunzioni, nato dallo studio delle oscillazioni, consiste nel trovare i valori λ per cui esiste una funzione non nulla f tale che

$$f(x) = \lambda \int_a^b k(x, y) f(y)\, dy.$$

Come prima, possiamo scrivere

$$f = \lambda A f$$

oppure

$$Af = \frac{1}{\lambda} f. \qquad (35)$$

Consideriamo ora A come un operatore lineare arbitrario. Un vettore non nullo f che soddisfi la (35) si chiama autovettore di A, e $1/\lambda$ è il corrispondente autovalore. Più in generale la definizione è $Af = \mu f$, ammettendo anche $\mu = 0$.

Poiché $(1/\lambda)f$ differisce da f soltanto per un fattore numerico, la ricerca degli autovettori si può descrivere come la ricerca delle rette vettoriali che la trasformazione A lascia invarianti.

Questo punto di vista unifica i problemi agli autovalori delle equazioni integrali, quando A è integrale, delle equazioni differenziali, quando è differenziale, e dell'algebra lineare, quando agisce in dimensione finita (capitoli VI e XVI). In tre dimensioni il problema compare nella ricerca degli assi principali di un ellissoide.

Per le equazioni integrali, proprietà importanti come la realtà dei valori caratteristici e l'ortogonalità delle autofunzioni associate a valori distinti derivano dalla simmetria del nucleo: $k(x, y) = k(y, x)$.

L'analogo operatoriale di questa proprietà è l'autoaggiuntezza.

Per un operatore limitato definito su tutto lo spazio di Hilbert reale, l'autoaggiuntezza equivale all'uguaglianza, per ogni coppia di elementi f_1, f_2,

$$(Af_1, f_2) = (f_1, Af_2)$$

dove (Af_1, f_2) è il prodotto scalare di Af_1 e f_2. Per operatori non limitati l'identità sul dominio esprime soltanto la simmetria: l'autoaggiuntezza richiede inoltre che operatore e aggiunto abbiano lo stesso dominio.

Nei problemi meccanici, l'autoaggiuntezza è spesso conseguenza della conservazione dell'energia, con opportune condizioni al contorno. Compare, per esempio, negli operatori che descrivono oscillazioni senza dissipazione.

Anche in meccanica quantistica gli operatori che rappresentano le osservabili sono autoaggiunti.

Verifichiamo l'autoaggiuntezza di un operatore integrale limitato con nucleo reale simmetrico. In questo caso Af_1 è la funzione $\int_a^b k(x, y) f_1(y)\, dy$. Il prodotto scalare con f_2 è quindi

$$(Af_1, f_2) = \int_a^b \int_a^b k(x, y) f_1(y) f_2(x)\, dy\, dx.$$

Analogamente,

$$(f_1, Af_2) = \int_a^b \int_a^b k(x, y) f_2(y) f_1(x)\, dy\, dx.$$

L'uguaglianza $(Af_1, f_2) = (f_1, Af_2)$ segue immediatamente dalla simmetria del nucleo, scambiando le variabili d'integrazione.

Gli operatori autoaggiunti possiedono numerose proprietà importanti per le applicazioni. I loro autovalori sono reali e gli autovettori associati ad autovalori distinti sono ortogonali.

Dimostriamo la seconda affermazione. Siano λ_1 e λ_2 due autovalori distinti di A, e f_1, f_2 autovettori corrispondenti. Ciò significa che

$$\begin{aligned} Af_1 &= \lambda_1 f_1, \\ Af_2 &= \lambda_2 f_2. \end{aligned} \qquad (36)$$

Moltiplichiamo scalarmente la prima equazione (36) per f_2 e la seconda per f_1:

$$\begin{aligned} (Af_1, f_2) &= \lambda_1 (f_1, f_2), \\ (Af_2, f_1) &= \lambda_2 (f_2, f_1). \end{aligned} \qquad (37)$$

Poiché A è autoaggiunto, nel caso reale abbiamo $(Af_1, f_2) = (Af_2, f_1)$. Sottraendo le due equazioni (37), otteniamo

$$0 = (\lambda_1 - \lambda_2)(f_1, f_2).$$

Essendo $\lambda_1 \neq \lambda_2$, segue $(f_1, f_2) = 0$: i due autovettori sono ortogonali.

Lo studio degli operatori autoaggiunti ha chiarito molti problemi concreti della teoria degli autovalori. Soffermiamoci sugli sviluppi in autofunzioni nel caso dello spettro continuo.

Per spiegare questo concetto torniamo alla corda vibrante. Per una corda di lunghezza l le frequenze proprie assumono i valori

$$a\frac{\pi}{l}, 2a\frac{\pi}{l}, \cdots, na\frac{\pi}{l}, \cdots.$$

Rappresentiamo questi valori sull'asse numerico. Aumentando l, la distanza fra valori adiacenti diminuisce e i punti si distribuiscono sempre più fittamente. Nel limite $l \to \infty$, corrispondente a una corda infinita, le frequenze possibili riempiono il semiasse $\lambda \geqslant 0$. Si parla allora di spettro continuo.

Per una corda finita, gli sviluppi in autofunzioni sono sviluppi trigonometrici. In una rappresentazione mediante seni e coseni di $n\pi x/l$ hanno la forma

$$a_0/2 + \sum_{n=1}^{\infty} [a_n \cos nx + b_n \sin nx]$$

Anche per una corda infinita possiamo sviluppare una funzione sufficientemente regolare mediante seni e coseni. Ma, poiché le frequenze variano con continuità, la serie è sostituita da un integrale di Fourier:

$$f(x) = \int_{-\infty}^{+\infty} [A(\lambda) \cos \lambda x + B(\lambda) \sin \lambda x]\, d\lambda.$$

Gli sviluppi in integrali di Fourier erano già ben noti e largamente usati nell'Ottocento per risolvere problemi di fisica matematica.

Nei casi più generali di spettro continuo,[14] però, molti problemi relativi agli sviluppi in autofunzioni rimasero a lungo poco chiari. Solo la teoria generale degli operatori autoaggiunti fornì il chiarimento necessario.

Ricordiamo un altro insieme di problemi classici affrontati mediante la teoria generale degli operatori: le oscillazioni con dissipazione di energia.

Anche qui si possono cercare modi liberi nella forma $u(x)\phi(t)$. Nei casi oscillatori smorzati considerati, invece del semplice cos ωt, il fattore temporale assume la forma $e^{-kt}\cos \omega t$, con $k > 0$. La soluzione è dunque $u(x)e^{-kt}\cos \omega t$. Ogni punto oscilla con frequenza angolare ω, ma l'ampiezza si attenua: il fattore e^{-kt} tende a zero per $t \to \infty$.

È comodo scrivere i modi propri in forma complessa, $u(x)e^{-i\lambda t}$: senza attrito λ è reale; nei modi smorzati è complesso, con parte immaginaria negativa nella convenzione adottata.

Le oscillazioni dissipative conducono ancora a problemi agli autovalori, ma per operatori non autoaggiunti. Una caratteristica è la comparsa di valori complessi che descrivono lo smorzamento.

Combinando metodi operatoriali e teoria delle funzioni analitiche, M. V. Keldyš studiò nel 1950-1951 classi di questi problemi, dimostrando, nelle ipotesi appropriate, teoremi di completezza dei sistemi di autofunzioni e funzioni associate.

Rapporti con altri rami della matematica e con la meccanica quantistica.

Abbiamo già ricordato l'impulso decisivo della meccanica quantistica. Come le esigenze della meccanica e della fisica classica orientarono il calcolo differenziale e integrale nel Settecento, così la fisica contemporanea, soprattutto la meccanica quantistica, esercitò una forte influenza sull'analisi funzionale. L'apparato matematico fondamentale della meccanica quantistica appartiene essenzialmente a questa disciplina. Possiamo soltanto accennare ai legami, perché una spiegazione dei fondamenti fisici oltrepasserebbe i limiti del libro.

In meccanica quantistica uno stato puro è rappresentato da un vettore normalizzato dello spazio di Hilbert, identificando i vettori che differiscono per un fattore di fase. Grandezze come energia, quantità di moto e momento angolare sono studiate mediante operatori autoaggiunti. Per esempio, i livelli energetici legati di un elettrone in un atomo si calcolano come autovalori dell'operatore energia. Le differenze fra livelli determinano, attraverso la relazione $\Delta E = h\nu$, le frequenze dei fotoni emessi e quindi la struttura dello spettro di radiazione. Gli stati corrispondenti sono descritti dalle autofunzioni dell'operatore energia.

I problemi quantistici richiedono spesso di calcolare autovalori di operatori, di solito differenziali. Nei casi complessi la soluzione esatta può essere praticamente irraggiungibile. Si usa allora la teoria delle perturbazioni, che permette, sotto opportune ipotesi, di ricavare dagli autovalori e dalle autofunzioni noti di un operatore autoaggiunto A quelli di un operatore A_1 poco diverso. All'epoca del testo, la fondazione matematica completa di questa teoria era indicata come un problema ancora aperto, interessante e importante.

Anche senza calcoli approssimati degli autovalori, un'indagine qualitativa può fornire molte informazioni. In meccanica quantistica essa si basa spesso sulle simmetrie: quelle dei cristalli, la simmetria sferica di un atomo, l'invarianza per rotazioni e così via. Poiché le simmetrie formano gruppi (capitolo XX), i metodi della teoria delle rappresentazioni dei gruppi permettono di rispondere a numerosi problemi senza calcoli espliciti, per esempio nella classificazione degli spettri atomici e nello studio delle trasformazioni nucleari. La meccanica quantistica usa dunque ampiamente gli operatori autoaggiunti e, nello stesso tempo, ne stimola la teoria proponendo problemi nuovi.

L'influenza della meccanica quantistica e lo sviluppo interno dell'analisi funzionale hanno fatto sì che, nei decenni cui si riferisce l'opera, problemi e metodi algebrici assumessero un ruolo sempre maggiore. Questa tendenza è paragonabile alla crescente importanza dei metodi algebrici nella fisica teorica rispetto ai metodi della fisica ottocentesca.

Concludiamo sottolineando ancora una volta che l'analisi funzionale è uno dei rami in rapido sviluppo della matematica dell'epoca. I rapporti con la fisica, le equazioni differenziali e il calcolo approssimato, insieme all'uso di metodi generali dell'algebra, della topologia e della teoria delle funzioni reali, ne fanno uno dei punti d'incontro fondamentali della matematica.

Letture consigliate

N. Dunford and J. T. Schwartz, Linear operators. I. General theory, Interscience, New York, 1958.

I. M. Gel'fand and Z. Ja. Šapiro, Representations of the rotation group of three-dimensional space and their applications, Amer. Math. Soc. Translations Series 2, vol. 2, 1956, 207-316.

A. N. Kolmogorov and S. V. Fomin, Elements of the theory of functions and functional analysis. Vol. 1, Metric and normed spaces. Vol. 2, Measure, Lebesgue integrals and Hilbert space. Graylock, New York, 1957/1961.

L. D. Landau and E. M. Lifšic, Course of theoretical physics. Vol. 3, Quantum mechanics, Pergamon, New York, 1958/1960.

F. Riesz and B. Sz.-Nagy, Functional analysis, Frederick Ungar, New York, 1955.

A. E. Taylor, Introduction to functional analysis, Wiley, New York, 1958.

Note al capitolo XIX

[1] Si veda il capitolo XVI.

[2] Il fatto che $|\cos \phi| \leqslant 1$ segue dalla disuguaglianza (1′).

[3] Un altro esempio è dato dalle oscillazioni elettriche in una catena di circuiti collegati (figura 6).

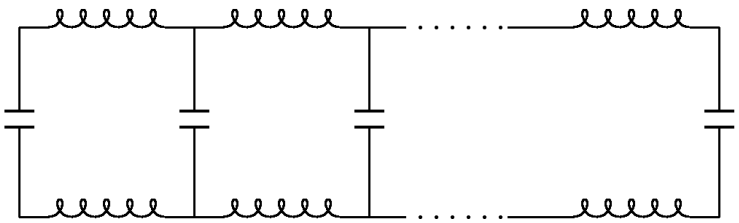


Figura 6

La configurazione elettrica della catena può essere descritta dai numeri $u_1, u_2, \ldots, u_n$, dove u_i è la tensione sul condensatore del circuito *i*-esimo. La successione $(u_1, \ldots, u_n)$ è un vettore dello spazio a *n* dimensioni.

Immaginiamo ora una linea a due fili, costituita da due conduttori con capacità e induttanza distribuite lungo la linea. La distribuzione della tensione è descritta da una funzione $u(x)$, vettore dello spazio a infinite dimensioni delle funzioni sull'intervallo (a, b).

[4] Si può prendere per *C* il valore assoluto dell'integrale, perché nella disuguaglianza quadratica si possono scegliere liberamente i segni dei parametri λ e μ. Se il prodotto integrale è negativo, si inverte opportunamente uno dei segni prima di usare la forma (9).

[5] Si veda il capitolo XII, §1.

[6] Sviluppi di questo tipo compaiono in molti problemi fisici quando si decompone un'oscillazione nelle sue componenti armoniche. Si veda il capitolo VI, §5.

[7] Per le serie trigonometriche si veda anche il capitolo XII, §7.

[8] La convergenza della serie alla somma è intesa nel senso precisato dalla formula (16).

[9] Gli altri integrali sono nulli perché le funzioni $\phi_k(x)$ sono a due a due ortogonali.

[10] Geometricamente, il quadrato della lunghezza di un vettore dello spazio di Hilbert è la somma dei quadrati delle sue proiezioni sulle direzioni di un sistema ortonormale completo.

[11] Questa affermazione sarà dimostrata nella sezione successiva.

[12] Se $k(x, y)$ è la funzione di Green di un sistema elastico, la condizione che l'equazione (29) abbia soltanto la soluzione nulla ha un significato fisico semplice. Abbiamo visto nella (23) che il carico distribuito $f(y)$ produce lo spostamento $u(x) = \int_a^b k(x, y) f(y)\, dy$. La condizione significa dunque che ogni carico non nullo nello spazio considerato sposta il sistema dalla configurazione di equilibrio.

[13] Si veda il capitolo VI, §5.

[14] Per esempio, le oscillazioni di un mezzo elastico non omogeneo e molti problemi di meccanica quantistica.

Nota editoriale

Sono conservate le cinque sezioni, le dieci figure numerate, le formule, la bibliografia e le quattordici note. La figura 6 è ridisegnata secondo lo schema originale, nel quale la didascalia si sovrapponeva al circuito. Le parole inglesi inserite in alcune formule sono tradotte; nella serie trigonometrica si uniforma il termine costante ad $a_0/2$, coerentemente con i coefficienti successivi.

Lo spazio di Hilbert delle funzioni a quadrato integrabile si intende completo, con l'integrale di Lebesgue e identificando le funzioni uguali quasi ovunque. Il quadrato della distanza non va confuso con la distanza stessa. Nella dimostrazione della disuguaglianza di Cauchy si separa il caso banale $C = 0$. Una famiglia ortonormale dà lo sviluppo di ogni elemento e l'identità di Parseval quando è completa.

Le affermazioni sulle equazioni integrali simmetriche si intendono sotto le ipotesi che rendono compatto e autoaggiunto l'operatore: per esempio, un nucleo reale simmetrico continuo su un quadrato compatto. I valori caratteristici non nulli dell'equazione $f = \lambda K f$ sono i reciproci degli autovalori non nulli di K. L'esistenza di infiniti autovalori non nulli richiede rango infinito. In presenza di molteplicità si sceglie una base ortonormale di ciascun autospazio.

Per gli operatori non limitati occorre specificare il dominio; la sola simmetria non equivale all'autoaggiuntezza. Gli operatori compatti mandano gli insiemi limitati in insiemi relativamente compatti. Le onde dello spettro continuo si intendono come modi generalizzati, non come vettori ordinari di L^2. Per la corda a estremi fissi bastano i modi sinusoidali; la serie con seni e coseni esprime il più generale sviluppo periodico.

Si distingue la pulsazione ω dalla frequenza ordinaria $\nu = \omega/(2\pi)$. Per le oscillazioni dissipative si intende il regime oscillatorio; nei risultati di completezza per problemi non autoaggiunti possono occorrere anche funzioni associate. Il richiamo quantistico alla frequenza del fotone usa la differenza tra le energie dei due stati. Le valutazioni sullo sviluppo della disciplina e i riferimenti a ricerche recenti rispecchiano l'epoca dell'opera.

CAPITOLO XX

Gruppi e altri sistemi algebrici

A. I. Mal'cev

§1. Introduzione

Nel capitolo IV, dedicato all'algebra dei polinomi, abbiamo già parlato delle principali linee di sviluppo dell'algebra, del suo posto fra le discipline matematiche e dei mutamenti nella concezione del suo stesso oggetto. Questo capitolo intende dare un'idea delle nuove teorie algebriche sorte nell'Ottocento, pienamente sviluppate nel Novecento e capaci di esercitare una profonda influenza sulla ricerca matematica contemporanea all'opera.

Come l'algebra classica, l'algebra moderna studia le operazioni e le regole di calcolo. Non si limita però alle proprietà delle operazioni sui numeri: cerca di studiare operazioni su elementi di natura assai più generale. Questa tendenza risponde a esigenze pratiche. In meccanica, per esempio, si sommano forze, velocità o rotazioni. Nell'algebra lineare (capitolo XVI), le cui idee e i cui metodi trovano larghe applicazioni nei calcoli pratici, si opera su matrici, trasformazioni lineari o vettori di uno spazio a n dimensioni.

La teoria dei gruppi occupa un posto particolarmente importante nell'algebra moderna, e le è dedicata gran parte di questo capitolo. Fra le altre teorie ci soffermeremo su quella dei sistemi ipercomplessi, tappa necessaria e importante nello sviluppo storico del concetto di numero. Queste due teorie non esauriscono affatto il contenuto dell'algebra moderna, ma ne illustrano bene le idee e i metodi.

La teoria dei gruppi è nata dall'esigenza di disporre di uno strumento per indagare regolarità fondamentali del mondo reale, come la simmetria.

La conoscenza delle proprietà di simmetria dei corpi geometrici o di altri oggetti matematici e fisici fornisce talvolta la chiave per chiarirne la struttura. Sebbene il concetto di simmetria sia del tutto intuitivo, una descrizione precisa e generale, e in particolare una caratterizzazione quantitativa delle sue proprietà, richiede gli strumenti della teoria dei gruppi.

La teoria dei gruppi nacque fra la fine del Settecento e l'inizio dell'Ottocento. In origine si sviluppò soltanto come strumento ausiliario per risolvere mediante radicali le equazioni di grado superiore. Proprio in questo problema si riconobbe per la prima volta che le proprietà di equivalenza e di simmetria delle radici sono decisive per la soluzione. Nell'Ottocento e nel Novecento il ruolo delle leggi di simmetria emerse in molti altri rami della scienza: geometria, cristallografia, fisica e chimica. Ne derivò una vasta diffusione dei metodi e dei risultati della teoria dei gruppi. Ogni campo di applicazione le proponeva problemi particolari; il moltiplicarsi delle applicazioni esercitò quindi anche un'influenza inversa, facendo nascere nuove branche della teoria. Pur restando unitaria nei concetti essenziali, essa si articolò così in

discipline relativamente autonome: teoria generale dei gruppi, teoria dei gruppi finiti, dei gruppi continui, dei gruppi discreti di trasformazioni, teoria delle rappresentazioni e dei caratteri, e così via. Nel corso di questa evoluzione i suoi metodi e concetti si rivelarono importanti non soltanto per studiare la simmetria, ma anche per risolvere molti altri problemi.

Il concetto di gruppo è divenuto uno dei concetti generali più importanti della matematica moderna, e la teoria dei gruppi ha assunto una posizione di primo piano fra le discipline matematiche. Contributi eminenti alla teoria e alle sue applicazioni furono dati da E. S. Fedorov, O. Ju. Šmidt e L. S. Pontrjagin. All'epoca dell'opera, le ricerche dei matematici sovietici occupavano un posto di rilievo nello sviluppo della disciplina.

§2. Simmetria e trasformazioni

Le forme più semplici di simmetria.

Cominciamo dalle forme più semplici di simmetria, familiari nella vita quotidiana. Una di esse è la simmetria speculare dei corpi geometrici, ossia la simmetria rispetto a un piano.

Un punto A dello spazio si dice simmetrico di un punto B rispetto a un piano α (figura 1) se il piano interseca perpendicolarmente il segmento AB nel suo punto medio. Si dice anche che B è l'immagine speculare di A nel piano α. Un corpo geometrico è simmetrico rispetto a un piano se questo lo divide in due parti, ciascuna immagine speculare dell'altra. Il piano si chiama allora piano di simmetria del corpo. La simmetria speculare si incontra spesso in natura: la forma del corpo umano, degli uccelli e di molti animali presenta abitualmente, in prima approssimazione, un piano di simmetria.

La simmetria rispetto a una retta si definisce analogamente. Due punti A e B sono simmetrici rispetto a una retta se questa interseca il segmento AB nel suo punto medio ed è perpendicolare ad AB (figura 2). Un corpo geometrico è simmetrico rispetto a una retta, o ha quella retta come asse di simmetria di ordine 2, se il simmetrico di ogni suo punto appartiene ancora al corpo.

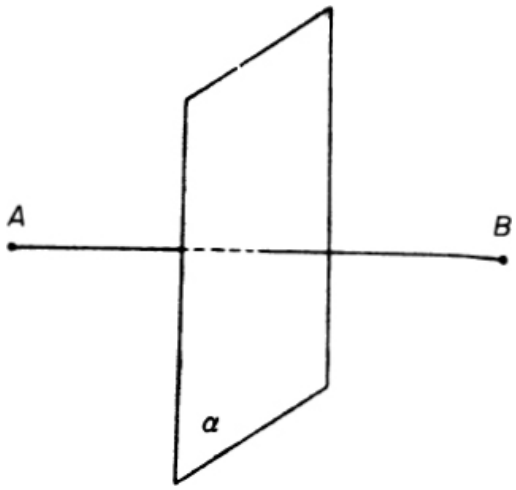


Figura 1.



Figura 2.

Un corpo che possiede un asse di simmetria di ordine 2 coincide con sé stesso quando viene ruotato attorno a quell'asse di mezzo giro, cioè di 180°.

Il concetto di asse di simmetria si generalizza naturalmente. Una retta è un asse di simmetria di ordine n per un corpo se il corpo coincide con sé stesso dopo una rotazione di $360°/n$ attorno alla retta. Per esempio, una piramide regolare avente per base un poligono regolare di n lati ha come asse di simmetria di ordine n la retta che congiunge il vertice al centro della base (figura 3).

Una retta si chiama asse di rivoluzione di un corpo se il corpo coincide con sé stesso per una rotazione di angolo arbitrario attorno a essa. Sono esempi l'asse di un cilindro o di un cono e qualunque diametro di una sfera. Un asse di rivoluzione è anche asse di simmetria di ogni ordine.

Un altro tipo importante è la simmetria rispetto a un punto, o simmetria centrale. I punti A e B sono simmetrici rispetto al centro O se O è il punto medio del segmento che li congiunge. Un corpo è simmetrico rispetto a O se il simmetrico di ogni suo punto appartiene al corpo. La sfera e il cubo sono esempi di corpi a simmetria centrale: i loro centri sono centri di simmetria (figura 4).

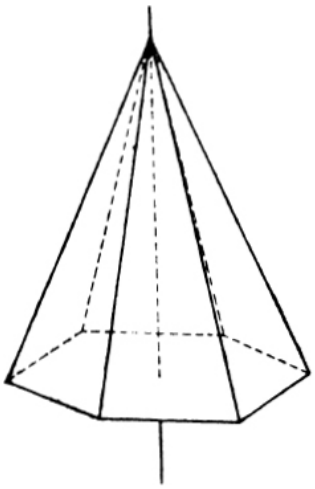


Figura 3.

La conoscenza di tutti i piani, gli assi e i centri di simmetria di un corpo dà un'idea abbastanza completa delle sue proprietà di simmetria.

Il concetto di simmetria non si applica soltanto alle figure geometriche. Ha un significato perfettamente chiaro dire che nel polinomio $x_1^3 + x_2^3 + x_3^3 + x_4^3$ le variabili x_1, x_2, x_3, x_4 compaiono simmetricamente; oppure che nel polinomio $x_1^3 + x_2^2 + x_3^2 + x_4^3$ compaiono simmetricamente x_1 e x_4, così come x_2 e x_3, mentre x_1 e x_2 hanno ruoli diversi. Gli esempi si potrebbero moltiplicare. Sorge così una domanda fondamentale: che cos'è, in generale, la simmetria, e come possiamo descriverla matematicamente? Una risposta precisa è legata al concetto di trasformazione, già incontrato molte volte fin dai primi capitoli. Per definire la simmetria in modo da comprendere casi tanto diversi quanto i corpi nello spazio e i polinomi, occorre formulare il concetto di trasformazione nella massima generalità.

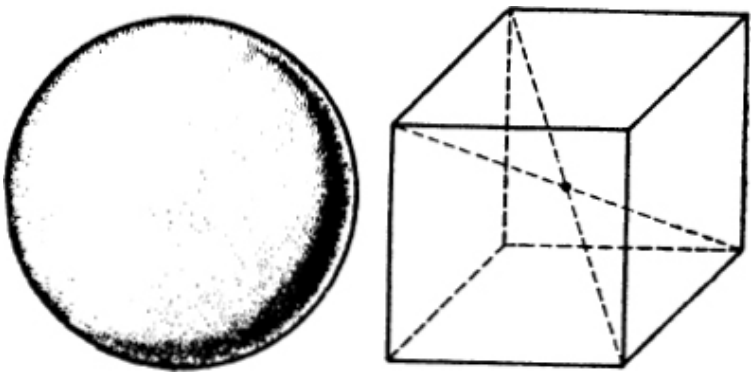


Figura 4.

Trasformazioni.

Sia M una collezione finita o infinita di oggetti del tutto arbitrari. Può essere, per esempio, l'insieme dei numeri 1, 2, ..., n, quello delle variabili indipendenti x_1, x_2, x_3, x_4, oppure quello di tutti i punti di un piano. Se a ogni elemento di M viene associato un elemento ben determinato dello stesso insieme, diciamo che è data una trasformazione di M. Per un insieme finito, ogni trasformazione può essere rappresentata da una tabella di due righe: nella prima si scrivono gli elementi in un ordine qualsiasi; sotto ciascuno si scrive l'elemento che gli corrisponde. Per esempio, la tabella

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 2 & 1 \end{pmatrix}$$

rappresenta la trasformazione dell'insieme 1, 2, 3, 4 che manda questi numeri rispettivamente in 2, 3, 2, 1. Scrivendo nella prima riga i numeri nell'ordine 3, 4, 1, 2, la stessa trasformazione assume la forma

$$\begin{pmatrix} 3 & 4 & 1 & 2 \\ 2 & 1 & 2 & 3 \end{pmatrix}.$$

Se M è infinito ma i suoi elementi sono numerabili, si può rappresentare la trasformazione allo stesso modo, disponendo gli elementi in un'unica successione; ciò avviene, per esempio, se M è l'insieme dei numeri naturali 1, 2, 3, ...

Per studiare le trasformazioni conviene introdurre una notazione generale. Le indicheremo semplicemente con lettere A, B, ecc. Se A è una trasformazione di M e m è un suo elemento, indicheremo con mA l'immagine di m, cioè l'elemento in cui esso viene mandato dalla trasformazione A. Supponiamo, per esempio, che

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 2 & 1 \end{pmatrix}$$

allora $1A = 2$, $2A = 3$, $3A = 2$, $4A = 1$.

Indichiamo alcune trasformazioni importanti in geometria.

Fissiamo nello spazio una retta arbitraria a e associamo a ogni punto P il punto Q ottenuto ruotando P attorno ad a di un angolo fissato ϕ (figura 5). Abbiamo così definito una trasformazione dell'insieme di tutti i punti dello spazio: la rotazione di angolo ϕ attorno all'asse a.

In meccanica la parola «rotazione» indica un processo attraverso il quale i punti di un corpo assumono nuove posizioni. Qui la usiamo nel senso di trasformazione dello spazio: prescindiamo dal processo effettivo del moto e consideriamo soltanto il risultato finale, cioè la corrispondenza fra le posizioni iniziali e quelle finali dei punti.

Un'altra trasformazione importante è la traslazione di tutti i punti dello spazio nella stessa direzione e per la stessa distanza. Nella figura 6 sono indicati, per alcuni punti arbitrari P_1, P_2,

P_3, i corrispondenti Q_1, Q_2, Q_3. È chiaro che, conoscendo l'immagine di un solo punto, si possono trovare quelle di tutti gli altri.

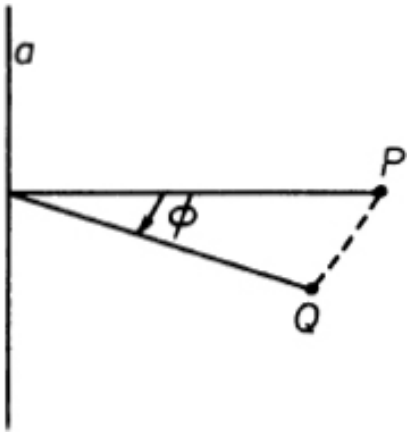


Figura 5.

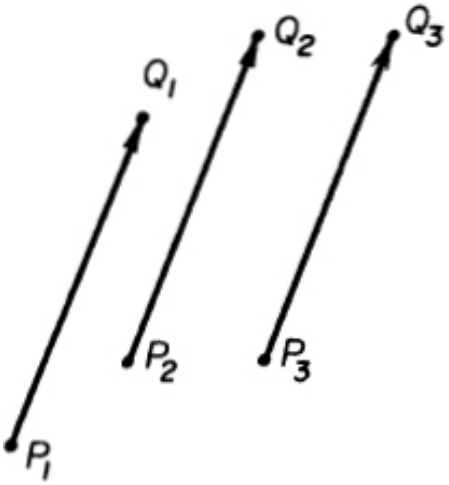


Figura 6.

Abbiamo definito il piano, l'asse e il centro di simmetria di una figura spaziale. A ciascuno corrisponde una trasformazione dello spazio: la riflessione nel piano, la rotazione di mezzo giro attorno alla retta e la simmetria centrale. Per esempio, nella riflessione in un piano ogni punto dello spazio viene associato al proprio simmetrico rispetto a quel piano. Analogamente si definiscono la simmetria assiale e quella centrale.

Finora abbiamo parlato di trasformazioni dello spazio. Nel piano si definiscono allo stesso modo la rotazione attorno a un punto di un angolo dato, la traslazione lungo una direzione data e la riflessione rispetto a una retta del piano. Queste trasformazioni sono ancora più intuitive delle corrispondenti trasformazioni spaziali.

Trasformazioni biunivoche.

Fra tutte le trasformazioni di un insieme bisogna distinguere quelle biunivoche dell'insieme su sé stesso dalle altre. Una trasformazione A di M è biunivoca se non soltanto a ogni elemento corrisponde un unico elemento di M, come richiede la definizione di trasformazione, ma anche ogni elemento y di M è immagine di uno e un solo elemento x. In altre parole, l'«equazione» $xA = y$ deve avere, per ogni y di M, una e una sola «soluzione» x in M.

Tutte le trasformazioni dello spazio considerate finora, riflessioni, rotazioni e traslazioni, sono biunivoche: ogni punto X ha un'immagine unica ed è a sua volta immagine di un unico punto.

È facile trovare esempi contrari. La trasformazione dell'insieme 1, 2, 3, 4 rappresentata dalla tabella

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \end{pmatrix}$$

non è biunivoca, perché nessun numero ha immagine 4. Anche la trasformazione dell'insieme di tutti i numeri naturali 1, 2, 3, ... rappresentata da

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & \cdots \\ 1 & 1 & 2 & 2 & 3 & 3 & \cdots \end{pmatrix}$$

non è biunivoca. Ogni numero n è immagine di $2n$, ma anche di $2n - 1$. Per le trasformazioni rappresentate da tabelle, il criterio di biunivocità è evidente: è necessario e sufficiente che nella seconda riga ogni elemento compaia una e una sola volta. In matematica si considerano spesso anche trasformazioni non biunivoche. Un esempio importante è la proiezione dello spazio su un piano: ogni punto del piano è immagine di un'intera retta di punti dello spazio. In molti casi, tuttavia, è conveniente trattare soltanto trasformazioni biunivoche. Esse hanno in particolare un ruolo fondamentale nella descrizione di processi fisici in cui gli elementi del sistema non si fondono, non vengono distrutti e non vengono creati.

D'ora in poi, parlando di trasformazioni, sottintenderemo che siano biunivoche. Esse si chiamano anche permutazioni, specialmente quando l'insieme è finito.

Per ogni trasformazione biunivoca A di M su sé stesso si definisce facilmente la trasformazione inversa A^{-1}. Se A manda un elemento arbitrario x in y, l'inversa manda y in x. Per esempio, se

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix},$$

allora

$$A^{-1} = \begin{pmatrix} 2 & 3 & 4 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix};$$

se A è una rotazione dello spazio attorno a un asse di angolo ϕ, A^{-1} è la rotazione dello stesso angolo attorno allo stesso asse nel verso opposto, e così via.

Talvolta l'inversa coincide con la trasformazione stessa. Ciò accade, in particolare, per le riflessioni rispetto a un piano e per le simmetrie centrali dello spazio. Accade anche per la permutazione

$$A = \begin{pmatrix} 2 & 1 & 4 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} \text{ ; poiché}$$

$$A^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 4 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

Per le trasformazioni non biunivoche non esiste invece una trasformazione inversa definita sull'intero insieme: un elemento può non avere alcuna controimmagine oppure averne più di una.

La definizione generale di simmetria.

In matematica e nelle sue applicazioni raramente occorre considerare tutte le trasformazioni di un insieme. Gli insiemi stessi, infatti, non sono di solito pensati come semplici collezioni di elementi del tutto privi di relazioni reciproche. Questo è naturale: gli insiemi matematici sono immagini astratte di collezioni reali, i cui elementi presentano innumerevoli relazioni fra loro e con ciò che sta al di fuori dell'insieme. È però conveniente prescindere da gran parte di tali legami, conservando quelli essenziali. Si è così indotti a considerare anzitutto le trasformazioni che rispettano le relazioni prescelte fra gli elementi. Esse si chiamano spesso trasformazioni ammissibili, o automorfismi rispetto a quelle relazioni. Per i punti dello spazio, per esempio, è importante la distanza: ogni coppia di punti ha una distanza determinata. Le trasformazioni che conservano questa relazione sono precisamente quelle che lasciano invariate le distanze. Si chiamano movimenti dello spazio, o isometrie.

Grazie al concetto di automorfismo possiamo dare una definizione generale di simmetria. Sia M un insieme nel quale si considerano determinate relazioni fra gli elementi, e sia P una sua parte. Diciamo che P è simmetrico, o invariante, rispetto a una trasformazione ammissibile A di M se A manda P su sé stesso, cioè $PA = P$. La simmetria di P è dunque caratterizzata dall'insieme delle trasformazioni ammissibili di M che conservano P. La simmetria di un corpo nello spazio rientra pienamente in questa definizione: M è l'intero spazio, le trasformazioni ammissibili sono i movimenti e P è il corpo. La sua simmetria è caratterizzata dai movimenti che lo fanno coincidere con sé stesso.

Riflessioni, traslazioni e rotazioni attorno a una retta sono casi particolari di movimenti, perché conservano le distanze. Un'indagine più approfondita mostra che ogni movimento del piano è una traslazione, una rotazione attorno a un punto, una riflessione in una retta oppure una riflessione seguita da una traslazione lungo quella retta, detta glissoriflessione. Analogamente, nello spazio ogni movimento è una traslazione, una rotazione attorno a un asse, un movimento elicoidale, cioè una rotazione combinata con una traslazione lungo lo stesso asse, oppure una riflessione in un piano eventualmente combinata con una traslazione parallela al piano o con una rotazione attorno a un asse perpendicolare al piano.

Traslazioni, rotazioni e movimenti elicoidali dello spazio si chiamano movimenti propri, o di prima specie. Gli altri, comprese le riflessioni, sono movimenti impropri, o di seconda specie. Nel piano sono di prima specie traslazioni e rotazioni; sono di seconda specie le riflessioni e le loro composizioni con rotazioni o traslazioni.

È facile immaginare come un movimento di prima specie si possa ottenere mediante un moto continuo dello spazio, o del piano, in sé stesso. Ciò non è possibile per i movimenti di seconda specie, a causa della riflessione speculare che entra nella loro costruzione.

Si dice spesso che il piano è simmetrico in tutte le sue parti, o che tutti i suoi punti sono equivalenti. Nel linguaggio preciso delle trasformazioni, questo significa che ogni punto può essere sovrapposto a qualunque altro mediante un opportuno movimento.

La definizione generale comprende anche le simmetrie dei corpi e delle figure già descritte. Un corpo simmetrico rispetto a un piano α coincide con sé stesso dopo la riflessione in α; un corpo simmetrico rispetto a un centro O coincide con sé stesso dopo la simmetria centrale di centro O. Il grado di simmetria di un corpo o di una figura spaziale è quindi caratterizzato dall'insieme di tutti i movimenti, propri e impropri, che lo fanno coincidere con sé stesso. Quanto più questo insieme è ricco e vario, tanto maggiore è la simmetria. Se contiene soltanto la trasformazione identica, il corpo può dirsi privo di simmetria.

La simmetria di un quadrato nel piano è caratterizzata dai movimenti del piano che lo fanno coincidere con sé stesso. Questi devono fissare anche l'intersezione delle diagonali, cioè il centro; sono quindi rotazioni attorno al centro oppure riflessioni in rette che lo attraversano. Dalla figura 7 si vede che il quadrato ABCD è invariante per rotazioni attorno a O di angoli multipli di 90° e per riflessioni nelle diagonali AC, BD e nelle rette KL, MN. Questi otto movimenti ne caratterizzano la simmetria.

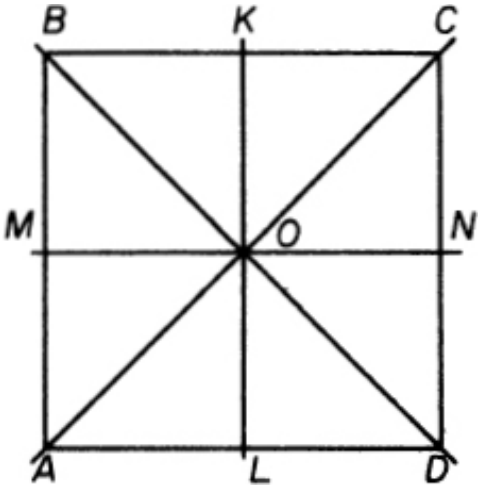


Figura 7.

Per un rettangolo non quadrato restano, oltre all'identità, la rotazione di 180° attorno al centro e le riflessioni nelle rette che congiungono i punti medi dei lati opposti. Per un parallelogramma generico (figura 8) restano soltanto le rotazioni di angoli multipli di 180°: l'identità e la simmetria centrale.

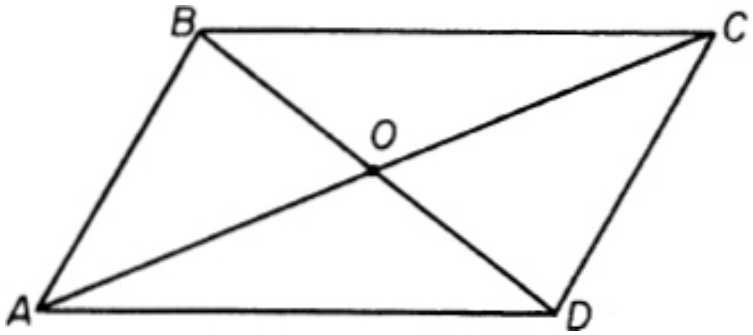


Figura 8.

Abbiamo già incontrato un esempio algebrico: anche per un polinomio in più variabili il concetto di simmetria ha un significato.

Vediamo come si può caratterizzare tale simmetria.

Diremo di aver eseguito nel polinomio la permutazione delle variabili

$$A = \begin{pmatrix} x_1 , x_2 , \cdots, x_n \\ x_{i_1} , x_{i_2} , \cdots, x_{i_n} \end{pmatrix}$$

o, più brevemente,

$$A = \begin{pmatrix} 1, 2, \cdots, n \\ i_1 , i_2 , \cdots, i_n \end{pmatrix}$$

se nel polinomio $F(x_1, x_2, ..., x_n)$ abbiamo sostituito ovunque x_1 con x_{i_1}, x_2 con x_{i_2}, e così via. Indicheremo con FA il polinomio ottenuto. Per esempio, se $F = x_1^2 - 2x_2 + x_3 - x_4^2$,

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

allora $FA = x_3^2 - 2x_1 + x_4 - x_2$.

La simmetria del polinomio è caratterizzata dall'insieme delle permutazioni delle variabili che lo lasciano invariato. Per esempio, la simmetria del polinomio $x_1^3 + 2x_2 + x_3^3 + 2x_4$ è caratterizzata dalle quattro permutazioni

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix},$$

mentre quella del polinomio $x_1^3 + 2x_2 + x_3^3 + x_4$ è caratterizzata dalle due permutazioni

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \text{ e } \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

§3. Gruppi di trasformazioni

Moltiplicazione delle trasformazioni.

Studiando le trasformazioni, si osserva facilmente che alcune si costruiscono a partire da altre. Un movimento elicoidale, per esempio, si compone di una rotazione attorno all'asse e di una traslazione lungo di esso. Questa costruzione si chiama moltiplicazione delle trasformazioni. Se a un elemento arbitrario x di M applichiamo prima A e poi applichiamo B all'elemento xA, otteniamo $(xA)B$. La trasformazione che manda direttamente x in $(xA)B$ si chiama prodotto di A e B e si indica con AB. Per definizione abbiamo quindi

$$x(AB) = (xA)B.$$

Esempio:

$$\begin{pmatrix}1&2&3&4\\2&3&4&1\end{pmatrix}\begin{pmatrix}1&2&3&4\\3&4&1&2\end{pmatrix}=\begin{pmatrix}1&2&3&4\\4&1&2&3\end{pmatrix}.$$

La prima permutazione manda 1 in 2 e la seconda manda 2 in 4; il prodotto deve dunque mandare 1 in 4, e così via. Ecco altri esempi:

$$\begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix}\begin{pmatrix}1&2&3&4\\1&3&2&4\end{pmatrix}^{-1}=\begin{pmatrix}1&2&3&4\\3&1&4&2\end{pmatrix}\begin{pmatrix}1&2&3&4\\1&3&2&4\end{pmatrix}=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix};$$

$$\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix}=\begin{pmatrix}1&2&3&4\\1&3&4&2\end{pmatrix};\begin{pmatrix}1&2&3&4\\3&1&2&4\end{pmatrix}\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}=\begin{pmatrix}1&2&3&4\\4&2&1&3\end{pmatrix}.$$

Gli ultimi due esempi mostrano che la moltiplicazione delle trasformazioni è, come si dice, non commutativa: il risultato dipende dall'ordine dei fattori. Lo si verifica facilmente anche per i movimenti del piano. Sia, per esempio, A la rotazione di 90° attorno all'origine O e sia B la traslazione di una unità lungo l'asse x.

Troviamo le immagini di O mediante AB e BA. Per definizione (figura 9),

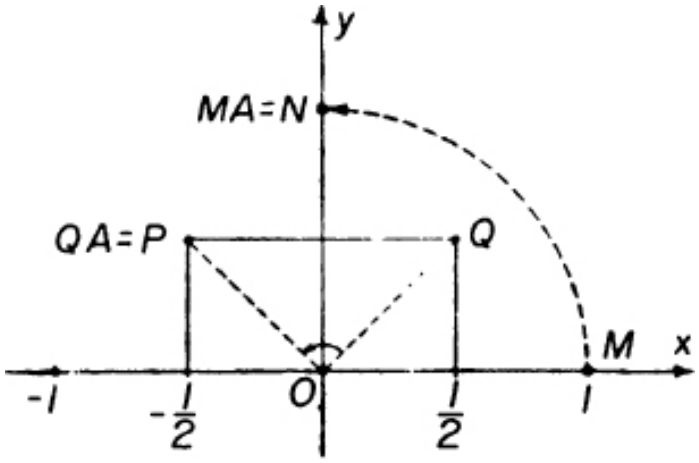


Figura 9.

$$O(AB) = (OA)\,B = OB = M,$$
$$O(BA) = (OB)\,A = MA = N,$$

dunque AB ≠ BA.

Per comprendere geometricamente BA consideriamo il punto *P*. Abbiamo

$$P(BA) = (PB)\,A = QA = P,$$

perciò *P* rimane fisso sotto BA. Partendo da questo fatto si mostra facilmente che BA è una rotazione del piano di 90° attorno a *P*. Analogamente,

$$Q(AB) = (QA)\,B = PB = Q,$$

e AB è la rotazione di 90° attorno a *Q*.

La moltiplicazione dei movimenti del piano o dello spazio segue in generale regole piuttosto complesse. In due casi importanti, però, le regole sono molto semplici. Il prodotto di due rotazioni del piano attorno allo stesso punto, o dello spazio attorno allo stesso asse orientato, di angoli ϕ e ψ, è la rotazione di angolo $\phi + \psi$. Inoltre, il prodotto delle traslazioni caratterizzate dai vettori $\overrightarrow{MN}$ e $\overrightarrow{NP}$ è la traslazione caratterizzata dal vettore $\overrightarrow{MP}$, somma dei vettori iniziali.

Il termine stesso «moltiplicazione» suggerisce un'analogia con quella dei numeri. L'analogia è però parziale: la moltiplicazione numerica è commutativa, mentre quella delle trasformazioni può non esserlo. La seconda legge fondamentale dell'aritmetica, l'associatività, rimane invece valida. Per trasformazioni arbitrarie *A*, *B*, *C* di *M* si ha *A*(BC) = (AB)*C*.

Infatti, per ogni elemento *m* di *M*,

$$m[A(BC)] = (mA)(BC) = [(mA)\,B]\,C = [m(AB)]\,C = m[(AB)\,C].$$

L'associatività permette di parlare di un unico prodotto *A*(BC) = (AB)*C* = ABC, senza distinguere le due parentesizzazioni. Anche il prodotto di quattro o più trasformazioni non dipende dalla disposizione delle parentesi.

Fra le trasformazioni ve n'è una che svolge il ruolo del numero 1: la trasformazione identica, o unitaria, *E*, che lascia fisso ogni elemento di *M*. Evidentemente AE = EA = *A* per qualsiasi *A*.

Il prodotto di trasformazioni biunivoche è ancora biunivoco. Per trovare l'elemento *x* mandato da AB in un elemento assegnato *a*, basta trovare l'unico x_1 mandato da *B* in *a* e poi l'unico x_2 mandato da *A* in x_1. Infatti x_2(AB) = (x_2A)*B* = x_1B = *a*, dunque x_2 è l'elemento cercato.

Il prodotto di una trasformazione *A* per la sua inversa A^{-1} è la trasformazione identica:

$$AA^{-1} = A^{-1}A = E.$$

Ciò segue immediatamente dalla definizione di trasformazione inversa.

L'esempio del prodotto di una traslazione e di una rotazione mostra che le proprietà del prodotto non si deducono sempre facilmente da quelle dei fattori. Un'importante eccezione è

la trasformazione $C = B^{-1}AB$: le sue proprietà sono legate in modo semplice a quelle di A e B. Se A manda m in n, allora C manda l'elemento «spostato» mB nell'elemento «spostato» nB.

Infatti: $(mB)B^{-1}AB = mAB = nB$.

Si dice che $B^{-1}AB$ si ottiene trasformando A mediante B, oppure che è la coniugata di A mediante B.

Coniughiamo, per esempio, una rotazione P_0 del piano attorno a O mediante una traslazione V. Per trovare le coppie di posizioni iniziali e finali del movimento $C = V^{-1}P_0V$, basta traslare mediante V le corrispondenti coppie di P_0. Poiché O rimane fisso sotto P_0 (figura 10), il punto OV rimane fisso sotto C. Se P_0 manda M in N, allora C manda MV in NV. Dalla figura si vede quindi che C è una rotazione attorno a OV dello stesso angolo ϕ.

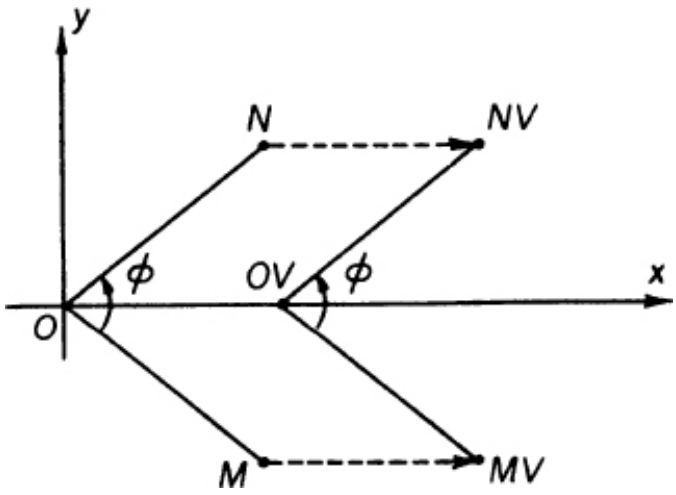


Figura 10.

Analogamente, coniugando una traslazione di vettore $\overrightarrow{MN}$ mediante una rotazione P_0 di angolo ϕ, si ottiene ancora una traslazione, il cui vettore è ruotato dello stesso angolo.

La regola per trovare $B^{-1}AB$ assume una forma particolarmente elegante quando le trasformazioni sono rappresentate da tabelle. Se

$$A = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & \cdots & n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix},$$

allora

$$B^{-1}AB = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ 1 & 2 & \cdots & n \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \cdots & n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ b_{a_1} & b_{a_2} & \cdots & b_{a_n} \end{pmatrix};$$

Per coniugare la permutazione A mediante B, basta dunque applicare B a tutti gli elementi della prima e della seconda riga di A. Per esempio, se

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 1 & 3 & 4 \end{pmatrix},$$

allora

$$B^{-1}AB = \begin{pmatrix} 1B & 2B & 3B & 4B & 5B \\ 3B & 5B & 4B & 1B & 2B \end{pmatrix} = \begin{pmatrix} 2 & 5 & 1 & 3 & 4 \\ 1 & 4 & 3 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 5 & 4 \end{pmatrix}.$$

Sebbene in generale il prodotto dipenda dall'ordine dei fattori, in certi casi AB e BA coincidono. Le trasformazioni *A* e *B* si dicono allora permutabili, o commutanti. Se AB = BA,

$$B^{-1}AB = B^{-1}BA = A.$$

La coniugazione mediante una trasformazione commutante lascia quindi invariata la trasformazione data.

Gruppi di trasformazioni.

L'insieme delle trasformazioni che caratterizzano la simmetria di una figura non può essere arbitrario: deve possedere le proprietà seguenti.

1. Il prodotto di due trasformazioni dell'insieme appartiene ancora all'insieme.

2. La trasformazione identica appartiene all'insieme.

3. Insieme a ogni trasformazione, l'insieme contiene anche la sua inversa.

Queste proprietà sono fondamentali. Ogni insieme di trasformazioni biunivoche di *M* su sé stesso che le soddisfa si chiama gruppo di trasformazioni di *M*, indipendentemente dal fatto che caratterizzi o meno la simmetria di una figura.

Dal punto di vista algebrico le proprietà 1-3 permettono di costruire, a partire da trasformazioni *A*, *B*, *C*, ... del gruppo, nuove trasformazioni come ABAC, $A^{-1}BCB^{-1}$, ecc., garantendo che esse appartengano ancora allo stesso insieme.

Il numero delle trasformazioni di un gruppo si chiama ordine del gruppo; può essere finito o infinito. Si distinguono così gruppi finiti e gruppi infiniti. Il gruppo di simmetria del quadrato nel piano ha otto elementi. Invece l'insieme infinito di punti A_i della figura 11 è conservato dalle traslazioni lungo OA, nei due versi, di distanze multiple di OA, dalle riflessioni nelle rette tratteggiate e dalla riflessione nell'asse OA, nonché dalle loro composizioni. Il suo gruppo di simmetria è dunque infinito.

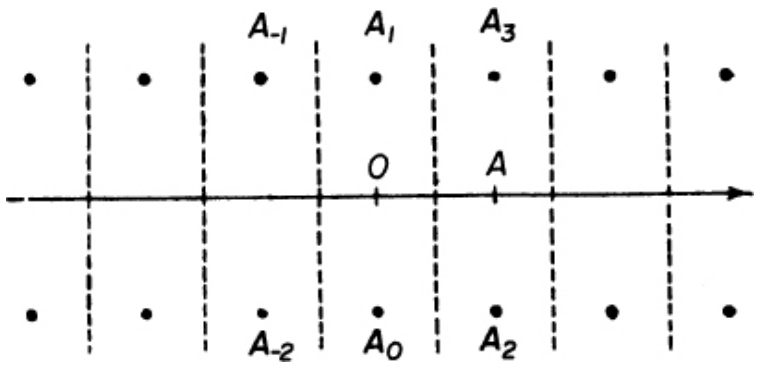


Figura 11.

Le trasformazioni che conservano un oggetto, caratterizzandone la simmetria, formano sempre un gruppo. Questo è uno dei modi più importanti di definire gruppi. Si ottengono così, anzitutto, i gruppi dei movimenti del piano e dello spazio. Sono molto interessanti anche i gruppi di simmetria dei poliedri regolari. Esistono cinque tipi di poliedri regolari convessi, con 4, 6, 8, 12 e 20 facce. Tutti i movimenti dello spazio che fanno coincidere un dato poliedro regolare con sé stesso costituiscono il suo gruppo di simmetria. Considerando soltanto i movimenti propri si ottiene ancora un gruppo, contenuto nel precedente, detto gruppo delle rotazioni del poliedro. Poiché ogni simmetria deve fissare il centro, i movimenti possibili sono rotazioni attorno ad assi passanti per il centro, riflessioni in piani passanti per il centro, oppure composizioni di tali riflessioni con rotazioni attorno ad assi perpendicolari ai piani.

Queste osservazioni permettono di determinare facilmente tutti i gruppi di simmetria e di rotazione dei poliedri regolari. La tabella 1 ne riporta gli ordini. Sono tutti gruppi finiti.

Tabella 1.

Numero delle facce	4	6	8	12	20
Ordine del gruppo di simmetria	24	48	48	120	120
Ordine del gruppo delle rotazioni	12	24	24	60	60

Gruppi di permutazioni.

Storicamente, i primi gruppi di trasformazioni studiati furono quelli ottenuti permutando le variabili x_1, x_2, ..., x_n nei polinomi. Il loro studio è strettamente legato alla risoluzione mediante radicali delle equazioni di grado superiore. Le permutazioni delle variabili che lasciano invariato uno o più polinomi formano evidentemente un gruppo. I polinomi invarianti per ogni permutazione delle variabili si chiamano simmetrici. È simmetrico, per esempio, $x_1 + x_2 + \ldots + x_n$. L'insieme di tutte le permutazioni delle variabili si chiama gruppo simmetrico.

Il numero delle variabili permutate è il grado del gruppo simmetrico. Al posto delle variabili possiamo permutare semplicemente i numeri 1, 2, ..., n. Ogni permutazione si scrive nella forma

$$\begin{pmatrix} 1\ 2 & \ldots\ n \\ a_1 a_2 & \ldots\ a_n \end{pmatrix},$$

dove a_1, a_2, ..., a_n sono i numeri 1, 2, ..., n disposti in un ordine qualsiasi. Il numero delle permutazioni, cioè l'ordine del gruppo simmetrico, è quindi $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$. Cresce molto rapidamente: per dieci elementi è già 3.628.800.

Consideriamo il polinomio

$$F(x_1, \cdots, x_n) = (x_2 - x_1)(x_3 - x_1) \cdots (x_n - x_1)(x_3 - x_2) \cdots (x_n - x_2) \cdots (x_n - x_{n-1}). \quad (1)$$

Ogni permutazione delle variabili lascia invariato F oppure ne cambia soltanto il segno. Nel primo caso la permutazione è pari, nel secondo dispari. Le permutazioni pari costituiscono il gruppo di simmetria del polinomio (1), detto gruppo alterno.

Il prodotto di due permutazioni pari è pari, perché esse formano un gruppo. Anche il prodotto di due permutazioni dispari è pari.

Infatti, se A e B sono dispari,

$$FAB = (FA)B = (-F)B = -(-F) = F.$$

Allo stesso modo si mostra che il prodotto di una permutazione pari e una dispari è dispari e che la permutazione inversa ha la stessa parità di quella iniziale.

Un esempio di permutazione dispari è

$$S = \begin{pmatrix} 1, 2, 3, \cdots, n \\ 2, 1, 3, \cdots, n \end{pmatrix},$$

che scambia gli elementi 1 e 2.

Decomposizione delle permutazioni in cicli.

Per studiare i gruppi di permutazioni è utile rappresentare ogni permutazione come prodotto di cicli. Il simbolo $(m_1, m_2, \ldots, m_k)$ indica la permutazione che manda m_1 in m_2, m_2 in m_3, …, m_{k-1} in m_k e infine m_k in m_1, lasciando fissi tutti gli altri elementi. Per le permutazioni di 1, 2, 3, 4, 5, per esempio,

$$(1, 2, 3, 4, 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}, \quad (3, 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 5 & 4 & 3 \end{pmatrix}.$$

Una permutazione di questa forma si chiama ciclo di lunghezza k; m_1, …, m_k ne sono gli elementi. L'identità si può scrivere mediante cicli di lunghezza 1: (1) = (2) = … I cicli di lunghezza 2 si chiamano trasposizioni. Spostando ciclicamente gli elementi nella scrittura di un ciclo si ottiene la stessa permutazione: (1, 2, 3) = (2, 3, 1) = (3, 1, 2), e (5, 6) = (6, 5).

Si verifica facilmente che cicli senza elementi comuni, come (2, 3) e (1, 4, 5), commutano; nel loro prodotto l'ordine dei fattori è quindi irrilevante.

L'importanza dei cicli si fonda sul teorema seguente: ogni permutazione di un insieme finito si rappresenta come prodotto di cicli disgiunti, e tale rappresentazione è unica a meno dell'ordine dei fattori e della scrittura ciclica di ciascuno di essi. I cicli di lunghezza 1 possono essere omessi.

La dimostrazione risulta dal procedimento stesso. Supponiamo di voler decomporre la permutazione

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 3 & 2 & 1 \end{pmatrix}.$$

La trasformazione A manda 1 in 4, 4 in 3, 3 in 6 e 6 in 1: il primo fattore è dunque (1, 4, 3, 6). Fra i numeri rimasti prendiamo 2: A manda 2 in 5 e 5 in 2, dando il secondo fattore (2, 5). Tutti i numeri sono ora esauriti, perciò

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 3 & 2 & 1 \end{pmatrix} = (1, 4, 3, 6)\,(2, 5). \qquad (2)$$

Si possono anche usare cicli con elementi comuni, ma la decomposizione non è più unica. Per esempio,

$$\begin{aligned} (a_1\,, a_2\,, \cdots, a_n) &= (a_1\,, a_2)(a_1\,, a_3) \cdots (a_1\,, a_n) \\ &= (a_2\,, a_3)(a_2\,, a_4) \cdots (a_2\,, a_n)(a_2\,, a_1). \qquad (3) \end{aligned}$$

Mostriamo che ogni ciclo di lunghezza 2 è dispari. Lo sappiamo già per (1, 2). Ogni ciclo (i, j) si scrive però $S^{-1}(1, 2)S$, dove S è una qualsiasi permutazione

$$\begin{pmatrix} 1, 2, \cdots \\ i\,, j\,, \cdots \end{pmatrix}$$

che manda 1 in i e 2 in j. Tale prodotto è dispari, perché (1, 2) è dispari e S, S^{-1} hanno la stessa parità.

Per la (3), un ciclo di lunghezza $m + 1$ è prodotto di m trasposizioni. È dunque dispari se la sua lunghezza è pari, e pari se la sua lunghezza è dispari. Ciò permette di calcolare rapidamente la parità di una permutazione di cui si conosca la decomposizione in cicli. In particolare, la permutazione

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 3 & 2 & 1 \end{pmatrix}$$

è pari, perché per la (2) è prodotto di due permutazioni dispari.

Sottogruppi.

Una parte di un gruppo che sia essa stessa un gruppo si chiama sottogruppo. Il gruppo alterno delle permutazioni di x_1, ..., x_n è un sottogruppo del gruppo simmetrico. I movimenti propri del piano formano un sottogruppo del gruppo di tutti i movimenti, propri e impropri.

La sola trasformazione identica costituisce un sottogruppo, e ogni gruppo è sottogruppo di sé stesso. Oltre a questi sottogruppi banali, i gruppi ne contengono generalmente molti altri. Conoscere tutti i sottogruppi dà un quadro abbastanza completo della struttura interna del gruppo.

Uno dei metodi più usati per costruire sottogruppi consiste nell'assegnarne i generatori.

Siano A_1, A_2, ..., A_m trasformazioni arbitrarie di un gruppo G. L'insieme H di tutti i prodotti finiti di queste trasformazioni e delle loro inverse è un gruppo. Contiene l'identità, che si scrive $A_1A_1^{-1}$. Se B e C sono prodotti di questo tipo, concatenando i loro fattori si ottiene una rappresentazione analoga di BC. Infine, se per esempio $B = A_1^{-1}A_2A_1A_1A_2^{-1}$, anche l'inversa ha la forma richiesta, poiché $B^{-1} = A_2A_1^{-1}A_1^{-1}A_2^{-1}A_1$.

Il gruppo H è un sottogruppo di G, detto sottogruppo generato da A_1, ..., A_m; queste trasformazioni sono i suoi generatori. Se H coincide con G, esse generano l'intero gruppo. È facile verificare con esempi che uno stesso sottogruppo può avere sistemi di generatori diversi.

Un sottogruppo generato da una sola trasformazione A si chiama ciclico. I suoi elementi sono le trasformazioni

$$E, A, AA, AAA, \cdots, A^{-1}, A^{-1}A^{-1}, A^{-1}A^{-1}A^{-1}, \cdots,$$

dette naturalmente potenze di A. Infatti,

$$E = A^0, A = A^1, AA = A^2, \cdots, A^{-1}A^{-1} = A^{-2},$$
$$A^{-1}A^{-1}A^{-1} = A^{-3}, \cdots.$$

Come nell'aritmetica ordinaria, si dimostra facilmente che

$$A^mA^n = A^{m+n} \text{ e } (A^m)^n = A^{mn}. \quad (4)$$

Una trasformazione è periodica se una sua potenza di esponente intero positivo è l'identità. Il minimo di tali esponenti si chiama ordine della trasformazione. Una trasformazione non periodica si dice di ordine infinito.

Sia A, per esempio, la rotazione del piano attorno a O di $360°/n$, con n intero maggiore di 1. Le potenze A^2, A^3, ..., A^{n-1} sono le rotazioni di $2(360°/n)$, $3(360°/n)$, ..., $(n-1)(360°/n)$, mentre A^n è la rotazione di 360°, cioè l'identità. La trasformazione ha dunque ordine n.

Se A è una traslazione non nulla lungo una retta, le potenze A^2, A^3, ... traslano lungo la stessa retta di distanze doppie, triple, ecc. Nessuna potenza positiva è l'identità: l'ordine è infinito.

Gli elementi del gruppo ciclico generato da A sono

$$\cdots, A^{-2}, A^{-1}, E, A, A^2, \cdots. \qquad (5)$$

Se A ha ordine infinito, tutti i termini della successione (5) sono distinti e il gruppo è infinito. Infatti, da $A^k = A^l$, con $k < l$, seguirebbe $A^{l-k} = E$ con esponente positivo, contro l'ipotesi.

Supponiamo ora che A abbia ordine finito m. Allora

$$A^m = E, A^{m+1} = A, A^{m+2} = A^2, \cdots, A^{m-1} = A^{-1}, A^{m-2} = A^{-2}, \cdots;$$

La successione (5) consiste dunque nei termini E, A, A^2, ..., A^{m-1}, ripetuti periodicamente. Questi termini sono distinti: se $A^k = A^l\ (0 \leqslant k < l < m)$, avremmo $A^{l-k} = E$ con $0 < l -$

$k < m$, contro la minimalità di m. Il gruppo ciclico generato da una trasformazione di ordine m contiene quindi esattamente m elementi.

Un gruppo i cui elementi commutano a due a due si chiama commutativo, o abeliano, in onore del matematico norvegese Abel, che ne riconobbe l'importanza per la teoria delle equazioni.

Per le formule (4), le potenze della stessa trasformazione commutano sempre: $A^m A^n = A^n A^m = A^{m+n}$. Ogni gruppo ciclico è quindi abeliano.

Nell'aritmetica, accanto alla moltiplicazione è importante la divisione. Poiché nei gruppi il prodotto può non essere commutativo, bisogna distinguere due divisioni. La soluzione di $Ax = B$ si chiama quoziente destro, quella di $yA = B$ quoziente sinistro di B per A. Moltiplicando la prima equazione a sinistra e la seconda a destra per A^{-1}, si ottengono $x = A^{-1}B$ e $y = BA^{-1}$. Entrambe le espressioni possono quindi essere considerate «quozienti» di B per A.

Abbiamo visto numerosi esempi in cui AB ≠ BA. Il «quoziente» $(AB)(BA)^{-1}$ oppure $(BA)^{-1}(AB)$ può servire a misurare la non commutatività. La seconda espressione, $(BA)^{-1}(AB) = A^{-1}B^{-1}AB$, si chiama commutatore di A e B e si indica con (A, B). Dalla formula

$$(A, B) = A^{-1}B^{-1}AB$$

segue che il commutatore è il quoziente della trasformazione coniugata $B^{-1}AB$ per A.

Se A è una traslazione del piano, ogni sua coniugata è ancora una traslazione; anche il quoziente di due traslazioni è una traslazione. Il commutatore di una traslazione con un movimento arbitrario del piano è quindi una traslazione. Sia invece A una rotazione di angolo ϕ attorno a O e sia B una rotazione o una traslazione. La coniugata è una rotazione dello stesso angolo attorno a un punto O'. Il commutatore è dunque il prodotto della rotazione di $-\phi$ attorno a O e di quella di ϕ attorno a O'. Dalla figura 12 si vede che il risultato è una traslazione di distanza $2 \cdot \overline{OO'} \sin(\phi/2)$, per l'angolo rappresentato, in direzione formante l'angolo $\pi/2 - \phi/2$ con il segmento $\overline{OO'}$.

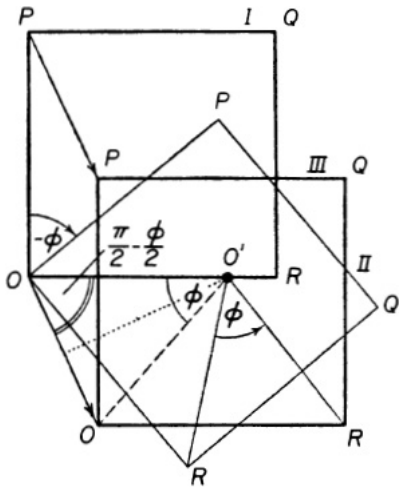


Figura 12.

Arriviamo così al fatto interessante che il commutatore di due movimenti propri del piano è sempre una traslazione, eventualmente nulla, cioè l'identità. Poiché $(A, B) = E$ equivale ad AB

= BA, ogni gruppo non commutativo di movimenti propri del piano contiene traslazioni non nulle.

Il sottogruppo generato da tutti i commutatori di elementi di *G* si chiama sottogruppo dei commutatori, o gruppo derivato di *G*. Esso consiste precisamente nei prodotti finiti di commutatori; anche l'inverso di un commutatore è un commutatore. Nel piano i commutatori di movimenti propri sono traslazioni e i prodotti di traslazioni sono ancora traslazioni: il gruppo derivato del gruppo dei movimenti propri contiene quindi soltanto traslazioni.

Il gruppo derivato di un gruppo abeliano contiene soltanto l'identità, perché AB = BA implica $(A, B) = E$.

Sia *G* il gruppo simmetrico su 1, 2, ..., *n*. Il commutatore di due permutazioni *A*, *B* è sempre pari. Infatti AB, BA e $(BA)^{-1}$ hanno la stessa parità; il loro quoziente $(A, B) = (BA)^{-1}(AB)$ è quindi pari.

Il gruppo derivato del gruppo simmetrico contiene soltanto permutazioni pari. Si dimostra facilmente che coincide con l'intero gruppo alterno.

Il derivato di *G* si indica spesso con *G′*; il suo derivato, detto secondo gruppo derivato di *G*, si indica con *G′′*. Ripetendo il procedimento si definiscono i gruppi derivati successivi.

Se uno dei gruppi derivati, e quindi tutti i successivi, contiene soltanto l'identità, *G* si dice risolubile. Il nome viene dalla teoria delle equazioni, dove la risolubilità del gruppo corrisponde alla risolubilità dell'equazione mediante radicali. Il gruppo dei movimenti propri del piano è risolubile, perché il secondo derivato è banale. I gruppi simmetrici di grado 2, 3 e 4 sono risolubili: rispettivamente il primo, il secondo e il terzo derivato sono banali. Quelli di grado almeno 5 non lo sono, perché il secondo derivato coincide con il primo e non è banale.

§4. Gruppi di Fedorov (gruppi cristallografici)

Gruppi di simmetria delle figure piane limitate.

La simmetria di una figura o di un corpo è caratterizzata dal gruppo dei movimenti del piano o dello spazio che lo fanno coincidere con sé stesso.

I gruppi più facili da determinare sono quelli delle figure piane limitate.[1] Se un movimento *A* conserva una tale figura, deve conservarne anche il centro di gravità *O*, quando questo è definito. Si tratta dunque di una rotazione attorno a *O* oppure di una riflessione in una retta passante per *O*. Il gruppo può contenere soltanto movimenti di questi tipi. Più in generale, per una figura limitata si può usare il centro del minimo disco che la contiene.

Esaminiamo i diversi casi possibili.

1. Il gruppo K_1 contiene soltanto l'identità. È il gruppo di una figura priva di simmetrie non banali (figura 13).

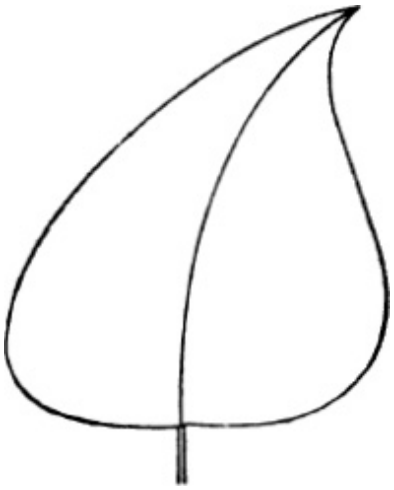


Figura 13.

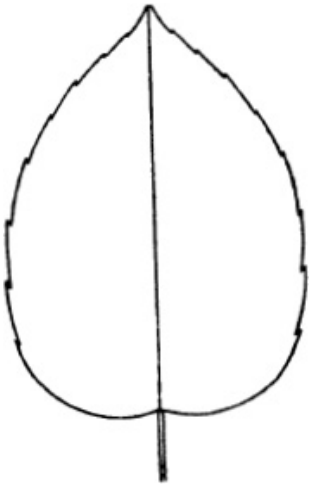


Figura 14.

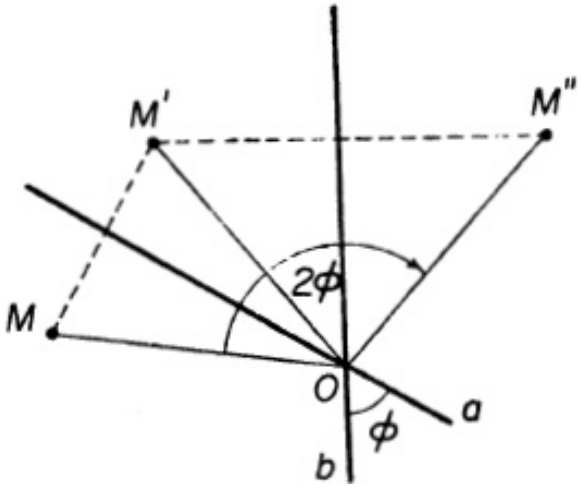


Figura 15.

2. Il gruppo K_2 contiene l'identità e la riflessione in una sola retta (figura 14).

Se un gruppo contiene le riflessioni in due rette a, b passanti per O e formanti un angolo ϕ, contiene anche il loro prodotto, cioè una rotazione attorno a O di angolo 2ϕ (figura 15). Pertanto K_2 è l'unico gruppo non banale della lista privo di rotazioni non identiche.

3. Il gruppo K_3 contiene soltanto rotazioni, ma nessuna di angolo positivo arbitrariamente piccolo. Esiste allora un minimo angolo positivo $\alpha°$. Ogni altro angolo del gruppo è un multiplo di $\alpha°$. Infatti, dato un angolo $\beta°$, scegliamo l'intero h per cui $h\alpha° \leqslant \beta° < (h+1)\alpha°$, così che $0 \leqslant \beta° - h\alpha° < \alpha°$. Il gruppo contiene anche la rotazione di $\beta° - h\alpha°$. Ma $0 \leqslant \beta° - h\alpha° < \alpha°$, e per minimalità di $\alpha°$ deve essere $\beta° - h\alpha° = 0$. Quindi $\beta° = h\alpha°$. Poiché il gruppo contiene la rotazione di 360°, si ha in particolare $n\alpha° = 360°$ per qualche intero n, ossia $\alpha° = 360°/n$.

Il gruppo K_3 consiste quindi nelle rotazioni di $0°$, $360°/n$, $2(360°/n)$, …, $(n-1)(360°/n)$. Facendo variare n fra 2, 3, 4, … si ottengono tutti i gruppi di questo tipo. La figura 16 mostra esempi per $n = 19$ e $n = 3$.

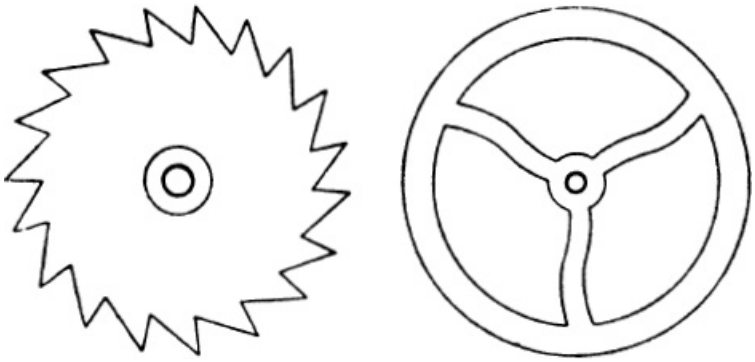


Figura 16.

4. Il gruppo K_4 contiene soltanto rotazioni, comprese rotazioni di angolo positivo arbitrariamente piccolo. Ogni angolo può allora essere approssimato quanto si vuole con angoli di rotazioni del gruppo. Consideriamo qui figure chiuse, comprendenti cioè i propri punti di frontiera (capitolo XVII, §9). Per tali figure il gruppo contiene tutte le rotazioni. È il caso della simmetria circolare orientata: una circonferenza, una corona circolare, ecc., dotata di un verso prescelto (figura 17). Le trasformazioni ammissibili devono conservare anche questo verso, e ciò esclude le riflessioni.

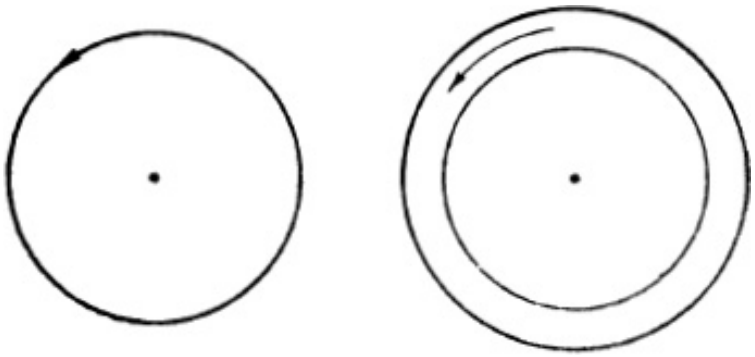


Figura 17.

Rimangono i casi misti, in cui il gruppo contiene sia rotazioni sia riflessioni. Senza svolgere la semplice dimostrazione, enunciamo il risultato: oltre ai tipi K_1-K_4, esistono soltanto i due seguenti.

5. Il gruppo K_5 contiene le riflessioni in n rette passanti per O, che dividono il piano in $2n$ angoli uguali, e le rotazioni di angoli multipli di $360°/n$. È, per esempio, il gruppo di simmetria di un poligono regolare di n lati (figura 18).

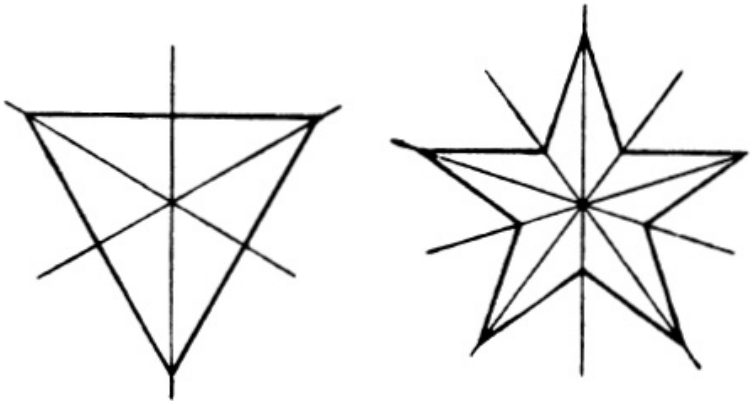


Figura 18.

6. Il gruppo K_6 contiene tutte le rotazioni attorno a O e tutte le riflessioni nelle rette passanti per O. È la simmetria circolare completa di una circonferenza o di una corona circolare non orientata.

Gruppi di simmetria delle figure piane illimitate.

Determinare tutti i gruppi di simmetria delle figure illimitate è più difficile. In pratica non disponiamo mai di un intero piano infinito, ma spesso una porzione di piano è coperta da figure tanto piccole da apparire, rispetto a esse, infinitamente estesa. Per esempio, la superficie piana e levigata di un pezzo d'acciaio presenta figure microscopiche, la cui regolarità rivela l'omogeneità della struttura interna del metallo.

Altri esempi sono i motivi ripetuti delle carte da parati e dei tessuti decorativi. L'arte dell'ornamento è diffusa presso moltissimi popoli dall'antichità fino a oggi. La figura 19 riproduce un motivo egizio per soffitti risalente alla metà del secondo millennio a.C.

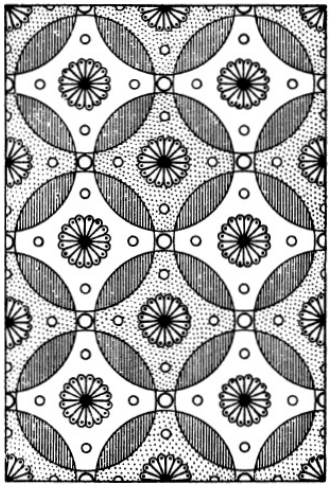


Figura 19.

Per le figure limitate abbiamo distinto i casi 1, 2, 3 e 5, privi di rotazioni arbitrariamente piccole, dai casi 4 e 6, che le contengono. Nello studio delle figure illimitate, soprattutto in tre dimensioni, la distinzione fra gruppi discreti e gruppi con trasformazioni arbitrariamente piccole diviene ancora più importante. Precisiamola.

Un gruppo di movimenti del piano si dice discreto se attorno a ogni punto si può tracciare un cerchio tale che ogni movimento del gruppo lasci fisso il punto oppure lo porti fuori da quel cerchio.

Si possono determinare tutti i gruppi discreti di movimenti del piano; ciascuno è il gruppo di simmetria di una figura piana. Si distinguono naturalmente tre tipi.

I. Esiste un punto fisso per tutte le trasformazioni del gruppo. Rientrano qui i tipi K_1, K_2, K_3 e K_5 della lista precedente.

II. Non esiste un punto fisso comune, ma esiste una retta mandata su sé stessa da tutte le trasformazioni. Essa si chiama asse del gruppo. Questi gruppi descrivono gli ornamenti disposti in una striscia infinita, o fregio. Sono esattamente sette.

1. Il gruppo L_1, formato dalle traslazioni lungo l'asse di distanze multiple di un segmento a.

2. Il gruppo L_2, ottenuto aggiungendo ai generatori di L_1 la rotazione di 180° attorno a un punto dell'asse.

3. Il gruppo L_3, ottenuto da L_1 aggiungendo una riflessione in una retta perpendicolare all'asse.

4. Il gruppo L_4, ottenuto da L_1 aggiungendo la riflessione nell'asse.

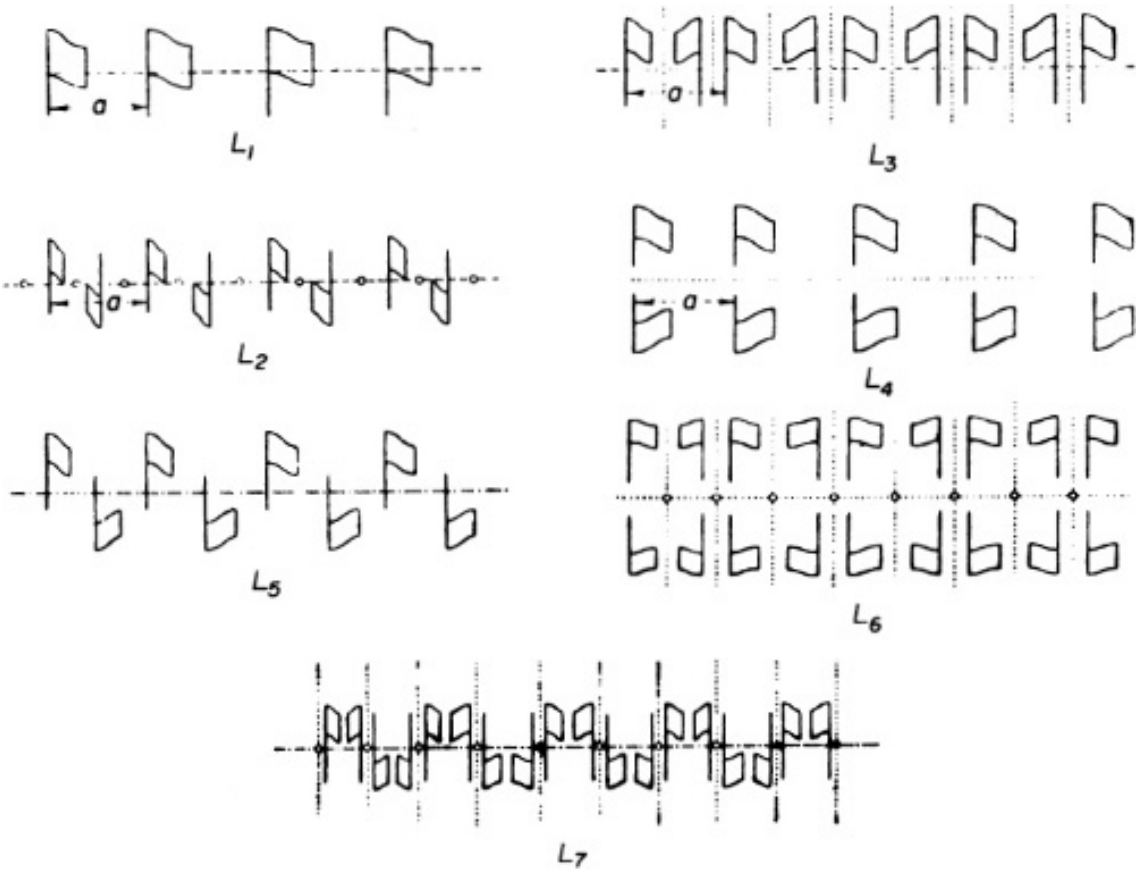
5. Il gruppo L_5, ottenuto da L_1 aggiungendo una traslazione di $a/2$ combinata con la riflessione nell'asse.

6. Il gruppo L_6, ottenuto da L_4 aggiungendo una riflessione in una retta perpendicolare all'asse.

7. Il gruppo L_7, ottenuto da L_5 aggiungendo una riflessione in una retta perpendicolare all'asse.

La tabella 2 presenta esempi di fregi corrispondenti ai gruppi L_1-L_7.

Tabella 2.



III. Non esistono né un punto né una retta invarianti per tutte le trasformazioni del gruppo. Questi sono i gruppi piani di Fedorov, gruppi di simmetria degli ornamenti estesi su tutto il piano. Sono diciassette: cinque contengono soltanto movimenti propri, dodici contengono anche movimenti impropri.

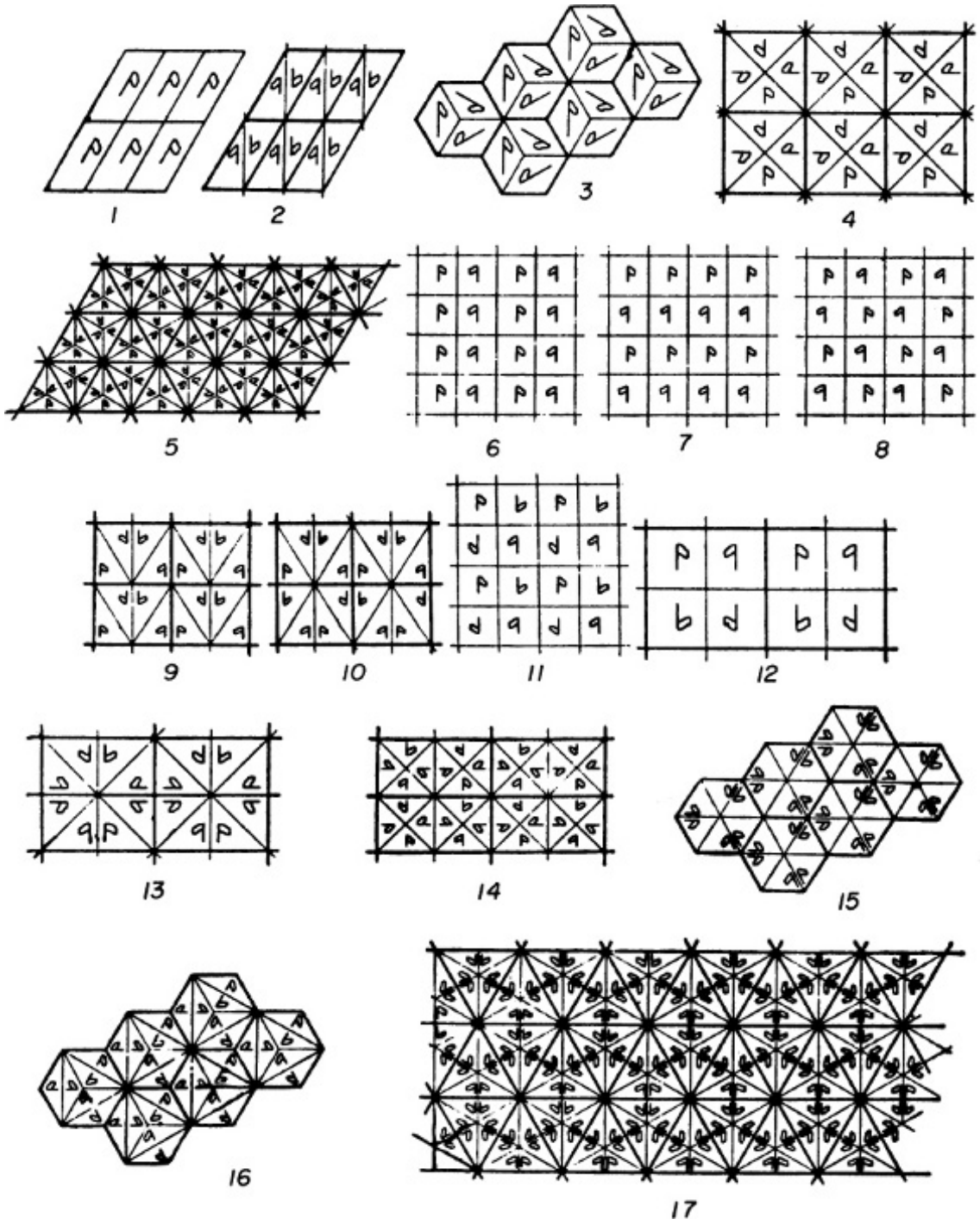
La tabella 3 presenta motivi corrispondenti a ciascuno dei diciassette gruppi. Ogni gruppo consiste nei movimenti che mandano una bandierina del disegno in un'altra conservando l'intera disposizione.

È interessante che gli artigiani dell'ornamento abbiano trovato nella pratica motivi con tutti i gruppi possibili: alla teoria dei gruppi spettava dimostrare che non ne esistono altri.

Gruppi cristallografici.

Nel 1890 il cristallografo e geometra russo E. S. Fedorov risolse con metodi della teoria dei gruppi un problema fondamentale della cristallografia: classificare i sistemi regolari di punti nello spazio. Fu un primo esempio importante di applicazione diretta della teoria a un problema delle scienze naturali, e influenzò considerevolmente il suo sviluppo.

Tabella 3.



In un cristallo gli atomi formano, in un senso da precisare, un sistema regolare nello spazio. Consideriamo i movimenti che mandano tale sistema su sé stesso: formano un gruppo, le cui proprietà permettono di definire con precisione la regolarità.

Un sistema di punti nello spazio si dice regolare se:

1. Ogni suo punto può essere mandato in qualunque altro mediante un movimento che conserva l'intero sistema.

2. Nessuna sfera di raggio finito contiene infiniti punti del sistema.

3. Esiste un numero positivo *r* tale che ogni sfera di raggio *r* contenga almeno un punto del sistema.

Lo studio della struttura dei cristalli è strettamente legato alla classificazione di questi sistemi e quindi dei gruppi discreti di movimenti dello spazio. Come nel piano, un gruppo *H* si dice discreto se, attorno a ogni punto *A*, esiste una sfera di raggio positivo tale che ogni movimento di *H* fissi *A* oppure lo porti fuori dalla sfera.

Si dimostra che i movimenti che conservano un sistema regolare di punti formano un gruppo discreto e che tutti i punti del sistema si ottengono da uno qualunque applicando tutte le trasformazioni del gruppo. Viceversa, da un gruppo discreto *H* e da un punto arbitrario *A* si ottiene, mediante tutti i movimenti di *H*, un sistema che soddisfa le condizioni 1 e 2. Condizioni aggiuntive permettono di individuare i gruppi per cui l'orbita soddisfa anche la condizione 3. Questi si chiamano gruppi di Fedorov, o gruppi cristallografici. La loro determinazione è dunque il primo passo fondamentale nello studio dei sistemi regolari. Per le applicazioni alle scienze naturali non basta considerare i gruppi di soli movimenti propri: occorrono anche quelli contenenti movimenti impropri, comprese le riflessioni. I primi sono molto meno numerosi; soltanto il caso generale permette di comprendere tutta la varietà delle strutture cristalline periodiche considerate dall'opera.

A differenza del caso piano, la grande quantità delle possibilità spaziali richiede gli strumenti della teoria dei gruppi per essere analizzata sistematicamente.

La maggiore complessità del problema spaziale rispetto a quello piano emerge dalla tabella 4.

Tabella 4. Numero dei gruppi spaziali di Fedorov

Gruppi con soli movimenti propri	65
Gruppi contenenti movimenti impropri	165
Totale	230

La deduzione e l'elencazione dettagliata di tutti i gruppi spaziali di Fedorov richiedono parecchie decine di pagine. Ci limitiamo dunque a questi risultati numerici, rinviando il lettore interessato alla letteratura specialistica.[2]

Gli sviluppi della cristallografia hanno richiesto ulteriori estensioni del concetto di simmetria. Nuove possibilità metodologiche sono delineate nel libro di A. V. Šubnikov, «Simmetria e antisimmetria delle figure finite», Accademia delle Scienze dell'URSS, Mosca, 1951.

§5. Gruppi di Galois

I risultati precedenti illustrano il ruolo dei gruppi nella classificazione dei cristalli. Questo problema non fu però all'origine della teoria. Circa un secolo prima, Lagrange aveva osservato un legame fra le simmetrie delle radici di un'equazione algebrica e la possibilità di risolverla mediante radicali. Nei primi decenni dell'Ottocento Abel e Galois approfondirono questo legame e giunsero a determinare le condizioni per la risolubilità mediante radicali. La soluzione si fondava su una sottile indagine delle proprietà dei gruppi di permutazioni e segnò l'inizio della teoria dei gruppi.

Lo studio dei legami fra equazioni algebriche e gruppi costituisce una vasta disciplina, chiamata teoria di Galois.

Nel capitolo IV abbiamo delineato la storia e il significato del problema. Poiché la teoria di Galois ebbe un ruolo decisivo nello sviluppo della teoria dei gruppi, ne riprendiamo i fatti fondamentali in una forma adatta a mettere in luce quest'ultima. Le dimostrazioni richiederebbero molti concetti ausiliari e saranno omesse.

Il gruppo di un'equazione algebrica.

Consideriamo un'equazione di grado n

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0, \qquad (6)$$

i cui coefficienti abbiano valori assegnati, per esempio numeri complessi. Il più piccolo campo contenente i numeri razionali e i coefficienti, ossia l'insieme delle quantità ottenibili da essi con un numero finito di addizioni, sottrazioni, moltiplicazioni e divisioni per quantità non nulle, si chiama campo di base, o dominio di razionalità dell'equazione.

Se i coefficienti sono razionali, il campo di base è il campo dei numeri razionali. Se l'equazione è $x^2 + \sqrt{2}\,x + 1 = 0$, esso è formato dai numeri $a + b\sqrt{2}$, con a e b razionali.

Indichiamo le radici con ξ_1, ..., ξ_n. Il campo ottenuto aggiungendole al campo di base e compiendo un numero finito delle quattro operazioni si chiama campo di spezzamento. Per $x^2 + 1 = 0$ esso è formato dai numeri complessi $a + bi$, con a, b razionali; per l'equazione $x^2 + \sqrt{2}\,x + 1 = 0$ è formato dai numeri $a + bi + c\sqrt{2} + di\sqrt{2}$, con a, b, c, d razionali.

Per le formule di Viète, i coefficienti dell'equazione monica si ottengono dalle radici mediante addizioni e moltiplicazioni. Il campo di spezzamento contiene quindi il campo di base. Talvolta i due coincidono.

Una corrispondenza biunivoca A del campo di spezzamento su sé stesso è un automorfismo relativo al campo di base se manda la somma di due elementi nella somma delle immagini, il prodotto nel prodotto delle immagini, e fissa ogni elemento del campo di base. In formule,

$$(a+b)A = aA + bA, \quad (ab)A = aA \cdot bA, \quad \alpha A = \alpha \qquad (a, b \in K, \quad \alpha \in P), \tag{7}$$

dove aA è l'immagine di a mediante A, P è il campo di base e K il campo di spezzamento.

Per il principio generale del §2, tutti questi automorfismi formano un gruppo, detto gruppo di Galois dell'equazione.

Per renderne più concreta l'idea, osserviamo che ogni automorfismo del gruppo manda una radice dell'equazione in un'altra radice. Se x soddisfa la (6), applicando A ai due membri e usando le proprietà (7) otteniamo

$$(xA)^n + (a_1A)(xA)^{n-1} + \ldots + a_nA = 0 \cdot A;$$

Poiché $0 \cdot A = 0$ e $a_iA = a_i$, segue

$$(xA)^n + a_1(xA)^{n-1} + \cdots + a_n = 0,$$

come volevamo. Ogni automorfismo induce dunque una permutazione delle radici. Viceversa, conoscendo tale permutazione si conosce l'automorfismo, perché ogni elemento del campo di spezzamento si esprime mediante operazioni aritmetiche a partire dalle radici e dal campo di base. Possiamo quindi considerare il corrispondente gruppo di permutazioni delle radici. In particolare, i gruppi di Galois delle equazioni qui considerate sono finiti.

Determinare il gruppo di Galois di un'equazione è in genere difficile; soltanto in casi particolari il compito è relativamente semplice. Consideriamo la (6) con coefficienti letterali indipendenti $a_1, \ldots, a_n$. Il campo di base è quello delle funzioni razionali dei coefficienti, cioè dei rapporti di polinomi in essi, a coefficienti razionali. Il campo di spezzamento è formato dalle funzioni razionali delle radici $\xi_1, \ldots, \xi_n$, legate ai coefficienti dalle formule

$$\begin{gathered} -a_1 = \xi_1 + \xi_2 + \ldots + \xi_n, \\ a_2 = \sum_{1 \le i < j \le n} \xi_i \xi_j, \\ \ldots \\ (-1)^n a_n = \xi_1 \xi_2 \ldots \xi_n. \end{gathered} \tag{8}$$

Poiché l'equazione è generale, possiamo trattare le radici come variabili indipendenti. Ogni loro permutazione determina allora un automorfismo del campo di spezzamento. Le formule (8) mostrano che questo fissa tutti i coefficienti, e quindi ogni funzione razionale di essi. Il gruppo di Galois dell'equazione generale di grado n è pertanto il gruppo simmetrico su n lettere.

Esistono anche equazioni a coefficienti numerici con gruppo di Galois simmetrico. Per esempio, si dimostra che l'equazione

$$1 - \frac{n}{1}x + \frac{n(n-1)}{1\cdot 2}\cdot\frac{1}{1\cdot 2}x^2 - \frac{n(n-1)(n-2)}{1\cdot 2\cdot 3}\cdot\frac{1}{1\cdot 2\cdot 3}x^3 + \cdots + (-1)^n\frac{1}{1\cdot 2\cdots n}x^n = 0 \qquad (9)$$

ha, per ogni grado n considerato, il gruppo simmetrico di grado n come gruppo di Galois.

Esistono metodi generali per costruire equazioni con un gruppo finito prescritto come gruppo di Galois, purché si possa scegliere liberamente il campo dei coefficienti. Se si richiedono coefficienti razionali, all'epoca dell'opera la costruzione era nota soltanto per particolari classi di gruppi. Un progresso notevole fu ottenuto da I. R. Šafarevič, che trovò metodi per realizzare ogni gruppo finito risolubile come gruppo di Galois di un'equazione a coefficienti razionali. Il problema generale era indicato dall'autore come ancora aperto.

Risolubilità delle equazioni mediante radicali.

Il gruppo di Galois caratterizza la simmetria interna delle radici. I problemi fondamentali sulla riduzione della soluzione di un'equazione a equazioni di grado inferiore, e molti altri problemi, si traducono in questioni sulla struttura di questo gruppo. Per un'equazione di grado n esso è un gruppo di permutazioni di grado n: un oggetto finito, nel quale tutte le relazioni possono essere determinate, almeno in linea teorica, per esame diretto.

Il gruppo di Galois è quindi uno strumento prezioso per le equazioni di grado superiore. In caratteristica zero, come nei casi qui trattati, un'equazione è risolubile mediante radicali se e soltanto se il suo gruppo di Galois è risolubile (si veda il §3). I gruppi simmetrici di grado 2, 3 e 4 sono risolubili, in accordo con le formule per le equazioni di questi gradi. I gruppi delle equazioni generali di grado 5, 6, ecc. sono invece i gruppi simmetrici dei rispettivi gradi, che non sono risolubili. Non esiste dunque una soluzione generale mediante radicali per i gradi superiori a 4.

Anche le equazioni (9), per $n > 4$, non sono risolubili mediante radicali, perché hanno gruppo di Galois simmetrico.

§6. Concetti fondamentali della teoria generale dei gruppi

Nell'Ottocento la teoria nacque principalmente come studio dei gruppi di trasformazioni. Divenne però sempre più chiaro che i risultati essenziali dipendevano soltanto dalla possibilità di moltiplicare le trasformazioni e da alcune proprietà caratteristiche del prodotto. Si incontrarono inoltre oggetti di altra natura, dotati di un'operazione con le stesse proprietà e ai quali si applicavano i medesimi teoremi. Alla fine dell'Ottocento il concetto di gruppo fu così esteso dai sistemi di trasformazioni a sistemi di elementi arbitrari.

Definizione generale di gruppo.

Supponiamo che a ogni coppia ordinata a, b di elementi di un insieme G venga associato un elemento ben determinato c dello stesso insieme. Diciamo allora che su G è definita un'operazione. Le si attribuiscono nomi come addizione, moltiplicazione o composizione; il risultato si chiama rispettivamente somma, prodotto o composto e si indica con $a + b$, ab o $a * b$. I nomi «addizione» e «moltiplicazione» si usano anche quando l'operazione non ha nulla a che fare con le consuete operazioni numeriche.

Un insieme G con un'operazione * si chiama gruppo rispetto a quell'operazione se valgono i seguenti assiomi.

1. Per ogni terna x, y, z di elementi di G,

$$x * (y * z) = (x * y) * z \text{ (legge associativa).}$$

2. Esiste in G un elemento e tale che, per ogni x,

$$x * e = e * x = x.$$

3. Per ogni elemento a esiste un elemento a^{-1} di G tale che

$$a * a^{-1} = a^{-1} * a = e.$$

L'elemento e dell'assioma 2 è l'elemento neutro; a^{-1} è l'inverso di a. Se l'operazione si chiama addizione, il neutro è detto zero e l'inverso opposto; se si chiama moltiplicazione, il neutro è l'unità. Gli assiomi assumono allora la forma

$$\begin{array}{ll} (1)\ x + (y + z) = (x + y) + z, & (1)\ x(yz) = (xy)z, \\ (2)\ x + 0 = 0 + x = x, & (2)\ xe = ex = x, \\ (3)\ x + (-x) = (-x) + x = 0, & (3)\ xx^{-1} = x^{-1}x = e. \end{array}$$

Nelle sezioni precedenti abbiamo visto molti gruppi di trasformazioni, con la composizione come operazione. Anche gli interi 0, ±1, ±2, … formano un gruppo rispetto all'addizione: la somma è ancora intera, l'operazione è associativa, lo zero è neutro e ogni a ha l'opposto $-a$. Un altro esempio è il gruppo moltiplicativo dei reali non nulli: il prodotto di due reali non

nulli è ancora non nullo, è associativo, 1 è neutro e ogni a ha l'inverso $a^{-1} = 1/a$. Gli esempi si potrebbero moltiplicare indefinitamente.

D'ora in poi chiameremo quasi sempre moltiplicazione l'operazione del gruppo. Sottogruppi, potenze, gruppi ciclici e ordine di un elemento si definiscono esattamente come per i gruppi di trasformazioni (§3). Ricordiamo soltanto che a e b sono coniugati se esiste x in G con $b = x^{-1}ax$. Ogni elemento è coniugato a sé stesso, poiché $a = a^{-1}aa$. Inoltre da $b = x^{-1}ax$ segue $a = xbx^{-1} = (x^{-1})^{-1}bx^{-1}$: la relazione è simmetrica. Infine, se $b = x^{-1}ax$ e $c = y^{-1}by$,

$$c = y^{-1}x^{-1}axy = (xy)^{-1}a(xy).$$

La relazione è quindi transitiva. Gli elementi del gruppo si suddividono così in classi disgiunte di elementi coniugati. Se il gruppo è commutativo, cioè $xy = yx$ per ogni coppia, due elementi coniugati coincidono: ogni classe contiene un solo elemento.

Isomorfismi.

Nel concetto di gruppo si distinguono due aspetti: la natura degli elementi e la legge con cui si moltiplicano. Si possono studiare i rapporti fra le proprietà particolari degli elementi, o dei loro insiemi, e quelle dell'operazione. Questo punto di vista è frequente per gruppi concreti, come quelli dei movimenti del piano o dello spazio. Si possono però studiare soltanto le proprietà esprimibili in termini dell'operazione: è il punto di vista della teoria astratta, o generale, dei gruppi. Il concetto di isomorfismo permette di precisarlo.

Due gruppi sono isomorfi se esiste una corrispondenza biunivoca fra i loro elementi che manda il prodotto di due elementi nel prodotto delle rispettive immagini. Tale corrispondenza si chiama isomorfismo.

Gli elementi corrispondenti hanno le stesse proprietà rispetto all'operazione. Un isomorfismo manda il neutro nel neutro, gli inversi negli inversi, gli elementi di ordine n in elementi dello stesso ordine e i sottogruppi in sottogruppi. La teoria astratta studia le proprietà conservate dagli isomorfismi. Per esempio, il gruppo delle permutazioni di quattro elementi e il gruppo completo di simmetria di un tetraedro regolare hanno, da questo punto di vista, le stesse proprietà. Ogni simmetria permuta i vertici; le simmetrie sono 24 e ogni permutazione dei quattro vertici si realizza in un unico modo. Associare a ciascun movimento la permutazione prodotta dà l'isomorfismo cercato.

Un esempio notevole viene dai logaritmi. Associando a ogni reale positivo il suo logaritmo, in una base fissata maggiore di 1, si ottiene una corrispondenza biunivoca con tutti i reali. La relazione $\log(xy) = \log x + \log y$ mostra che è un isomorfismo fra il gruppo moltiplicativo dei reali positivi e quello additivo di tutti i reali. La sua importanza pratica è ben nota.

Gruppi finiti di ordini diversi non sono isomorfi. Un gruppo astratto è determinato dalla legge di moltiplicazione, indipendentemente dalla natura degli elementi: gruppi concreti diversi ma isomorfi sono modelli dello stesso gruppo astratto.

Un gruppo astratto può essere assegnato in vari modi. Per i gruppi finiti il più immediato è la tavola di moltiplicazione.

Per un gruppo di ordine n, elencati gli elementi in un ordine arbitrario, si costruisce una tabella di n righe e n colonne. Nella casella di riga i e colonna j si scrive il prodotto dell'elemento i-esimo per il j-esimo. Questa tabella si chiama anche tavola di Cayley.

In pratica, tuttavia, la tavola completa è quasi sempre troppo ingombrante per essere un modo conveniente di assegnare il gruppo.

Vi sono altri metodi, fra cui l'assegnazione mediante generatori e relazioni definenti, già incontrata. Più spesso si specifica un gruppo concreto isomorfo a quello astratto, in particolare un gruppo di trasformazioni.

Ci si chiede allora se ogni gruppo astratto possa essere rappresentato come gruppo di trasformazioni. La risposta è il seguente teorema: ogni gruppo G è isomorfo a un gruppo di trasformazioni dell'insieme dei suoi stessi elementi.

Fissiamo g in G e indichiamo con A_g la trasformazione che manda ogni x in xg. È biunivoca, perché l'equazione

$$xA_g = xg = a$$

ha, per ogni a, l'unica soluzione $x = ag^{-1}$. Inoltre al prodotto gh corrisponde il prodotto A_gA_h, poiché

$$xA_{gh} = x(gh) = (xg)h = (xA_g)A_h = x(A_gA_h).$$

Al neutro corrisponde l'identità e all'inverso di g la trasformazione inversa. Le trasformazioni così ottenute formano quindi un gruppo Γ isomorfo a G. Se G ha più di due elementi, Γ non esaurisce tutte le permutazioni dell'insieme: è un sottogruppo proprio del gruppo simmetrico su quell'insieme.

Sottogruppi normali e gruppi quoziente.

Siano P e Q due sottoinsiemi di un gruppo G. Il prodotto PQ è l'insieme di tutti gli elementi che si scrivono come prodotto di un elemento di P per uno di Q. In particolare, gP consiste nei prodotti di g per tutti gli elementi di P.

Un sottogruppo H di G è normale, o invariante, se $gH = Hg$ per ogni g. Per un sottogruppo arbitrario, gli insiemi gH e Hg si chiamano rispettivamente laterale sinistro e laterale destro contenenti g. La normalità significa dunque che i due laterali corrispondenti allo stesso elemento coincidono.

Se H è normale, il prodotto di due laterali è ancora un laterale: $aH \cdot bH = abH$. Il sottogruppo H è esso stesso il laterale del neutro, o di qualunque suo elemento h, perché $hH = H$. Il prodotto dei laterali è associativo:

$$(aH \cdot bH)cH = (ab \cdot c)H = (a \cdot bc)H = aH(bH \cdot cH).$$

Il neutro è H, perché $H \cdot aH = eH \cdot aH = (ea)H = aH$, e analogamente $aH \cdot H = aH$. L'inverso di aH è $a^{-1}H$, dato che il loro prodotto è $aa^{-1}H = H$. Considerando i laterali come elementi di

un nuovo insieme, otteniamo quindi un gruppo, detto gruppo quoziente di G rispetto al sottogruppo normale H, e indicato con G/H.

Ogni laterale di un sottogruppo H di un gruppo finito contiene tanti elementi quanti H, e due laterali distinti sono disgiunti. Il numero dei laterali è dunque l'ordine di G diviso per quello di H. Ne segue il teorema di Lagrange: l'ordine di ogni sottogruppo di un gruppo finito divide l'ordine del gruppo.

In un gruppo abeliano ogni sottogruppo è normale. All'estremo opposto vi sono i gruppi semplici, non banali, che non hanno sottogruppi normali diversi dal gruppo stesso e dal sottogruppo del neutro. Sono importanti anche i gruppi risolubili del §3: essi ammettono una catena finita G, G_1, ..., G_k, nella quale ogni termine è un sottogruppo normale del precedente, l'ultimo è banale e tutti i quozienti G/G_1, G_1/G_2, ..., G_{k-1}/G_k sono abeliani.

Omomorfismi.

Il gruppo quoziente è strettamente legato al concetto di omomorfismo, fondamentale per tutta la teoria dei gruppi.

Un'applicazione di un gruppo G su un gruppo H è un omomorfismo se manda il prodotto di due elementi nel prodotto delle loro immagini. Qui, seguendo l'esposizione, consideriamo applicazioni suriettive; più in generale un omomorfismo può avere come immagine un sottogruppo di H.

Se indichiamo con x' l'immagine di x, la proprietà caratteristica è

$$(x_1x_2)' = x_1'x_2' .$$

Un isomorfismo è necessariamente biunivoco. Un omomorfismo, invece, assegna a ogni elemento di G un'unica immagine, ma elementi distinti possono avere la stessa immagine. Si può dire che nell'isomorfismo H è una copia esatta di G, mentre nell'omomorfismo più elementi possono «fondersi» in uno solo. Questa perdita di distinzione non è un difetto: rende gli omomorfismi strumenti potenti per lo studio dei gruppi.

Gli omomorfismi compaiono naturalmente nello studio delle trasformazioni. Consideriamo il gruppo di simmetria del tetraedro regolare (figura 20). È isomorfo al gruppo simmetrico su quattro elementi: ogni permutazione dei vertici A_1, A_2, A_3, A_4 è realizzata da un unico movimento proprio o improprio.

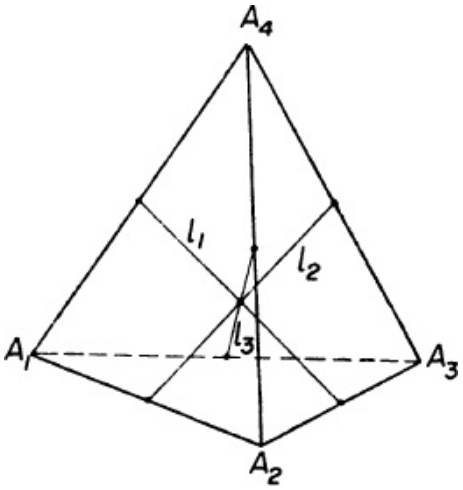


Figura 20.

Consideriamo ora le rette l_1, l_2, l_3 che congiungono i punti medi di spigoli opposti. Ogni simmetria del tetraedro le permuta, e ogni loro permutazione è indotta da qualche simmetria. Al prodotto di simmetrie corrisponde il prodotto delle permutazioni delle rette. La figura 20 illustra così un omomorfismo naturale dal gruppo simmetrico sui quattro vertici a quello sulle tre rette. Non è difficile individuare quali elementi del gruppo maggiore vengono identificati.

Un altro esempio: le permutazioni di n elementi formano un gruppo non commutativo se $n > 2$, mentre $\{+1, -1\}$ è un gruppo moltiplicativo. Associamo $+1$ alle permutazioni pari e -1 alle dispari. Otteniamo un omomorfismo, suriettivo per $n \geq 2$, perché il prodotto di permutazioni della stessa parità è pari e quello di parità diversa è dispari.

Associando a ogni reale non nullo x il valore assoluto $|x|$ si ottiene un omomorfismo dal gruppo moltiplicativo dei reali non nulli a quello dei reali positivi, poiché $|xy| = |x|\,|y|$.

Ogni movimento proprio A del piano si scrive come prodotto di una rotazione V_A attorno a un punto fissato O e di una traslazione D_A. Le rotazioni attorno a O formano un gruppo. La corrispondenza $A \to V_A$ manda il gruppo dei movimenti propri su quello delle rotazioni. È un omomorfismo: dalle decomposizioni $A = V_A D_A$ e $B = V_B D_B$ segue

$$AB = V_A D_A V_B D_B = (V_A V_B)\,(V_B^{-1} D_A V_B D_B).$$

La prima parentesi è una rotazione attorno a O; la seconda è il prodotto della traslazione coniugata $V_B^{-1} D_A V_B$ per D_B, dunque è ancora una traslazione. Al prodotto AB corrisponde quindi $V_A V_B$, come richiesto.

Infine, il gruppo quoziente G/N, per ogni sottogruppo normale N, è un'immagine omomorfa di G.

Basta associare a ogni g il laterale gN. Il prodotto gh viene mandato in $ghN = gN \cdot hN$, perciò l'applicazione è un omomorfismo suriettivo.

Passiamo alle proprietà generali: ogni omomorfismo manda il neutro nel neutro e gli inversi negli inversi.

Se e è il neutro di G ed e' la sua immagine, da $ee = e$ segue $e'e' = e'$. Indicando con ε il neutro di H e cancellando un fattore e', otteniamo $e' = \varepsilon$. Se poi x e y sono inversi, da $xy = e$ segue $x'y' = e' = \varepsilon$, dunque le immagini sono inverse e

$$(x^{-1})' = x'^{-1}.$$

Questi fatti permettono di trovare l'immagine di qualunque prodotto. Per esempio,

$$(ab^{-1}c^{-1}dh^{-1})' = a'(b^{-1})'(c^{-1})'d'(h^{-1})' = a'b'^{-1}c'^{-1}d'h'^{-1}.$$

Il teorema seguente è fondamentale nella teoria degli omomorfismi.

Dato un omomorfismo di G su H, l'insieme N degli elementi mandati nel neutro di H è un sottogruppo normale di G. Gli elementi che hanno una stessa immagine formano un laterale di N. La corrispondenza biunivoca così stabilita fra i laterali e gli elementi di H è un isomorfismo fra G/N e H.

Dimostriamolo. Per a, b appartenenti a N si ha $a' = b' = e'$, dove l'apice indica l'immagine in H. Quindi

$$(ab)' = a'b' = e'e' = e',$$
$$(a^{-1})' = a'^{-1} = e'^{-1} = e';$$

e sia il prodotto sia gli inversi appartengono a N, che è dunque un sottogruppo. Per ogni g di G si ha inoltre

$$(g^{-1}ag)' = g'^{-1}a'g' = g'^{-1}e'g' = g'^{-1}g' = e';$$

perciò $g^{-1}ag$ appartiene a N per ogni a di N. Questo implica la normalità e dimostra la prima affermazione.

Fissiamo ora g e sia U l'insieme degli elementi u con $u' = g'$. Se u appartiene a gN, si scrive $u = gn$ con n in N; allora $u' = g'n' = g'$, dunque gN è contenuto in U. Viceversa, se $u' = g'$, l'immagine di $g^{-1}u$ è $g'^{-1}u' = e'$. Quindi $g^{-1}u = n$ appartiene a N, e $u = gn$. Le due inclusioni danno $U = gN$.

Infine, ai laterali gN e hN corrispondono g' e h'. Al loro prodotto, per la formula

$$gN \cdot hN = ghN,$$

corrisponde $(gh)' = g'h'$, come richiesto.

Il teorema mostra che ogni immagine omomorfa H di G è isomorfa a G/N, dove N è il nucleo dell'omomorfismo. A meno di isomorfismo, tutte le immagini omomorfe si ottengono dunque come gruppi quoziente.

§7. Gruppi continui

Gruppi di Lie e gruppi continui di trasformazioni.

I successi ottenuti nella risoluzione delle equazioni algebriche spinsero i matematici, verso la metà dell'Ottocento, a cercare applicazioni della teoria dei gruppi ad altri tipi di equazioni, anzitutto a quelle differenziali, tanto importanti nelle applicazioni. Il tentativo riuscì. Benché i gruppi svolgano nelle equazioni differenziali un ruolo diverso, queste ricerche ampliarono profondamente il concetto di gruppo e diedero origine alla teoria dei gruppi continui e dei gruppi di Lie, fondamentale in molti rami della matematica.

I gruppi delle equazioni algebriche sono finiti; quelli incontrati nello studio delle equazioni differenziali possono essere infiniti. Nei casi considerati da Lie, tuttavia, le trasformazioni sono descritte localmente da un numero finito di parametri. Supponiamo che siano $a_1, \ldots, a_r$. Ai valori $x_1, \ldots, x_r$ corrisponde una trasformazione X; ai valori $y_1, \ldots, y_r$ una trasformazione Y. Il prodotto $Z = XY$ appartiene ancora al gruppo ed è descritto da parametri $z_1, \ldots, z_r$, funzioni dei parametri dei fattori:

$$
\begin{aligned}
z_1 &= \phi_1(x_1, x_2, \cdots, x_r;\ y_1, y_2, \cdots, y_r),\\
z_2 &= \phi_2(x_1, x_2, \cdots, x_r;\ y_1, y_2, \cdots, y_r),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
z_r &= \phi_r(x_1, x_2, \cdots, x_r;\ y_1, y_2, \cdots, y_r).
\end{aligned}
$$

I gruppi descritti localmente da un numero finito di parametri continui, con legge di moltiplicazione espressa da funzioni $\phi_1, \ldots, \phi_r$ sufficientemente differenziabili, sono detti gruppi di Lie, in onore del matematico norvegese Sophus Lie che per primo li studiò. Nella formulazione dell'opera si richiede la differenziabilità fino al secondo ordine; la definizione moderna usa una struttura di varietà differenziabile e richiede regolarità sia al prodotto sia all'inversione.

Nella prima metà dell'Ottocento N. I. Lobačevskij elaborò la geometria che porta il suo nome. Quasi contemporaneamente si affermò la geometria proiettiva e, poco più tardi, nacque quella di Riemann. Nella seconda metà del secolo esistevano così diversi sistemi geometrici, che studiavano da punti di vista differenti le «forme spaziali del mondo reale», secondo l'espressione di Engels. La teoria dei gruppi offrì un modo di riunire molte di queste geometrie in una prospettiva comune, conservandone le differenze essenziali.

Consideriamo le trasformazioni biunivoche dei punti di uno spazio geometrico che conservano le relazioni fondamentali studiate nella geometria in questione. Esse formano il gruppo dei movimenti, o degli automorfismi. Nel quadro delle geometrie omogenee considerate da Klein, la geometria può essere descritta come studio delle proprietà invarianti per quel gruppo. Il metodo di classificare le geometrie mediante i loro gruppi fu introdotto da F. Klein nella seconda metà dell'Ottocento ed è discusso nel capitolo XVII. I gruppi delle

principali geometrie classiche così studiate risultavano essere gruppi di Lie; ciò rese particolarmente importante la loro indagine.

Grazie ai legami con molti settori della matematica e della meccanica, la teoria dei gruppi di Lie si sviluppò intensamente fin dall'origine. Alcuni problemi vi furono risolti più rapidamente che per i gruppi finiti. L'autore osserva che, all'epoca, la classificazione dei gruppi finiti semplici, privi cioè di sottogruppi normali non banali, era ancora poco avanzata, mentre Killing e Cartan avevano ottenuto alla fine dell'Ottocento la corrispondente classificazione infinitesimale dei gruppi di Lie semplici. Ricorda inoltre i contributi di V. V. Morozov, A. I. Mal'cev ed E. B. Dynkin alla classificazione dei sottogruppi semplici dei gruppi di Lie. In un'altra direzione I. M. Gel'fand e M. A. Najmark svilupparono lo studio delle rappresentazioni continue mediante trasformazioni unitarie di uno spazio di Hilbert, di particolare interesse per l'analisi e la fisica.

Lo studio dei gruppi di Lie si avvale di uno strumento particolare: i cosiddetti gruppi infinitesimali, o algebre di Lie, che esamineremo nel §13.

Gruppi topologici.

Accanto allo sviluppo della teoria classica dei gruppi di Lie, si ottennero nell'URSS importanti progressi nella teoria più generale dei gruppi topologici, o continui. Il concetto è più ampio: non richiede un numero finito di parametri né leggi differenziabili. Un gruppo è topologico se è dotato di una topologia rispetto alla quale il prodotto e l'inversione sono continui. Intuitivamente, elementi vicini devono avere prodotti e inversi vicini.

Il concetto servì dapprima a organizzare idee fondamentali della teoria dei gruppi di Lie; in seguito ne emerse l'importanza per altri rami della matematica. I primi lavori sui gruppi topologici generali risalgono ai primi anni Venti del Novecento; i risultati che ne fecero una disciplina autonoma furono ottenuti soprattutto fra la fine degli anni Venti e l'inizio degli anni Trenta. Molti si devono a L. S. Pontrjagin, considerato uno dei fondatori della teoria moderna dei gruppi continui. Il suo libro «Gruppi topologici», primo trattato sistematico della materia, era ancora, a vent'anni dalla pubblicazione e all'epoca dell'opera, un riferimento fondamentale.

§8. Gruppi fondamentali

Negli esempi precedenti i gruppi sono comparsi quasi sempre come gruppi di trasformazioni. Facevano eccezione i gruppi numerici additivi e moltiplicativi. Esaminiamo ora un esempio importante in cui il gruppo nasce direttamente come sistema algebrico con un'operazione.

Il gruppo fondamentale.

Su una superficie S facciamo muovere un punto M lungo una curva continua da A a B: otteniamo un cammino. Esso può intersecare sé stesso e ripercorrere più volte alcuni tratti. Non basta perciò conoscere la curva tracciata: occorre sapere quante volte e in quale verso viene percorsa. Una stessa circonferenza, percorsa un numero diverso di volte o in versi diversi, dà cammini distinti. Due cammini con gli stessi estremi sono equivalenti se si trasformano l'uno nell'altro mediante una deformazione continua che mantiene fissi gli estremi; si dicono omotopi a estremi fissi. Nel piano e sulla sfera tutti i cammini da A a B sono equivalenti (figura 21). Sul toro, invece, i cammini chiusi U e V con punto iniziale e finale A nella figura 22 non lo sono.

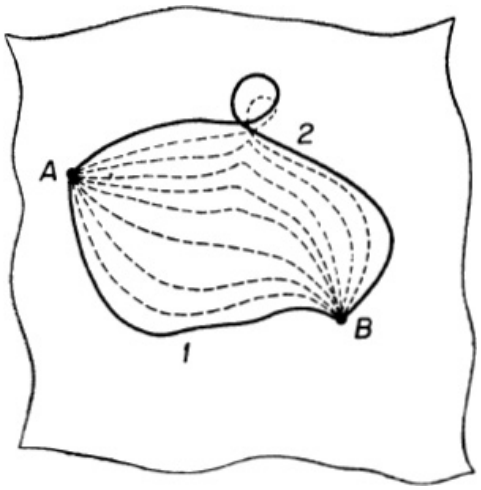


Figura 21.

Su un cilindro circolare infinito nei due versi, scegliamo il cammino X della figura 23. Ogni cammino chiuso che parte da A è equivalente a X^n, con n intero. Per n positivo si percorre X n volte; X^0 è il cammino costante in A; X^{-n} percorre X n volte nel verso opposto. Nella figura, $Z \sim X^{-1}$, $Y \sim X^2$, $U \sim X^0$. L'enorme varietà dei cammini si riduce così, a meno di equivalenza, al numero di giri orientati attorno al cilindro. Potenze con esponenti diversi non sono equivalenti.

Su una superficie arbitraria, siano U un cammino da A a B e V uno da B a C. Percorrendo prima U e poi V si ottiene il cammino prodotto UV da A a C. Sostituendo i fattori con cammini equivalenti, il prodotto rimane equivalente. La moltiplicazione è associativa a meno di equivalenza: U(VW) e (UV)W, quando definiti, rappresentano cammini equivalenti. Il

cammino inverso U^{-1} percorre U nel verso opposto, da B ad A. Il prodotto di un cammino per il proprio inverso è equivalente al cammino costante in A.

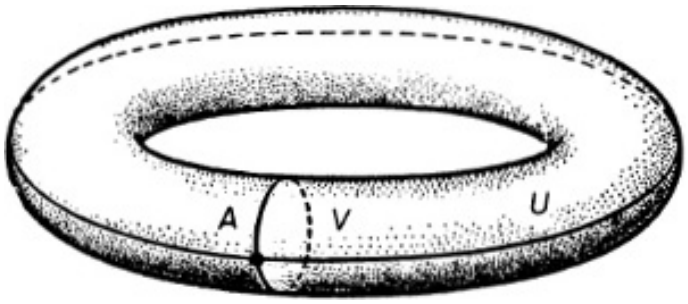


Figura 22.

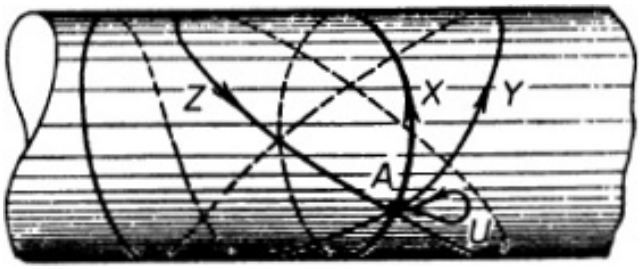


Figura 23.

Il prodotto è definito soltanto quando la fine del primo cammino coincide con l'inizio del secondo. Questa limitazione scompare se consideriamo esclusivamente cammini chiusi con lo stesso punto base A: il prodotto e l'inverso sono ancora cammini chiusi basati in A.

Consideriamo ora i cammini equivalenti come rappresentazioni di un unico elemento. Le osservazioni precedenti mostrano che le classi di equivalenza dei cammini chiusi basati in A formano un gruppo. Il neutro è la classe del cammino costante e l'inverso è rappresentato dal cammino percorso nel verso opposto.

Il gruppo può dipendere dalla superficie e dal punto base. Se però due punti della superficie si possono sempre congiungere mediante un cammino, i gruppi ottenuti scegliendo punti base diversi sono isomorfi. A meno di isomorfismo possiamo allora parlare semplicemente del gruppo dei cammini della superficie S, detto anche gruppo fondamentale.

Per il piano e per la sfera il gruppo è banale, perché ogni cammino chiuso si contrae a un punto. Per la superficie del cilindro infinito è invece ciclico infinito: ogni classe è una potenza di X, e potenze diverse sono distinte. Per il toro (figura 22) si dimostra che ogni elemento si scrive $U^m V^n$, con m, n interi, che UV = VU e che due espressioni di questa forma sono uguali soltanto se i rispettivi esponenti coincidono. Qui l'uguaglianza significa sempre equivalenza dei cammini.

L'importanza del gruppo fondamentale deriva dalla sua invarianza topologica. Supponiamo che fra S e un'altra superficie S_1 esista una corrispondenza biunivoca continua con inversa continua, cioè un omeomorfismo. Una deformazione senza strappi né identificazioni di punti ne offre un'immagine intuitiva. A ogni cammino di S corrisponde un cammino di S_1;

l'equivalenza e il prodotto sono conservati. I gruppi fondamentali sono quindi isomorfi. Se due superfici hanno gruppi non isomorfi, non possono essere omeomorfe. Il piano, per esempio, non può essere trasformato in un cilindro infinito senza tagli o identificazioni, perché il suo gruppo è banale e quello del cilindro è infinito.

Le proprietà conservate dagli omeomorfismi sono studiate dalla topologia, introdotta nel capitolo XVIII, e si chiamano invarianti topologici. Il gruppo fondamentale è uno degli esempi più importanti. Si definisce non soltanto per superfici, ma per qualunque spazio nel quale abbiano senso i cammini continui e le loro deformazioni.

Relazioni definenti.

La topologia studia metodi dettagliati per calcolare i gruppi fondamentali. È spesso conveniente assegnarli mediante un procedimento usato anche nella teoria generale dei gruppi: generatori e relazioni.

Gli elementi $g_1, \ldots, g_n$ generano G se ogni elemento si scrive nella forma

$$g = g_{i_1}^{\alpha_1} g_{i_2}^{\alpha_2} \cdots g_{i_k}^{\alpha_k},$$

dove gli indici $i_1, \ldots, i_k$ appartengono a 1, ..., n, il numero dei fattori è arbitrario e gli esponenti sono interi positivi o negativi. Indici non consecutivi possono coincidere; fattori consecutivi con lo stesso indice possono essere riuniti.

Per conoscere il gruppo basta sapere, oltre ai generatori, quali prodotti rappresentano lo stesso elemento e quali elementi distinti. Occorre dunque conoscere tutte le uguaglianze della forma

$$g_{i_1}^{\alpha_1} g_{i_2}^{\alpha_2} \cdots g_{i_k}^{\alpha_k} = g_{j_1}^{\beta_1} g_{j_2}^{\beta_2} \cdots g_{j_l}^{\beta_l}$$

valide nel gruppo. Poiché sono infinite, si assegnano normalmente soltanto relazioni dalle quali tutte le altre seguano usando gli assiomi di gruppo. Sono le relazioni definenti.

Uno stesso gruppo può essere presentato in modi diversi mediante generatori e relazioni.

Consideriamo, per esempio, il gruppo H con generatori a, b e relazioni

$$a^2 = b^3, \; ab = ba. \qquad (10)$$

Ponendo $c = ab^{-1}$, otteniamo

$$a = bc, \; a^2 = b^2c^2, \; b^3 = b^2c^2, \; b = c^2, \; a = c^3.$$

Tutti gli elementi si esprimono dunque mediante il solo c, con

$$a = c^3, \; b = c^2.$$

Le relazioni (10) seguono da queste uguaglianze senza imporre alcuna relazione non banale a c. Il gruppo H è perciò ciclico infinito, generato da c.

Un gruppo si dice libero se ammette generatori fra i quali non valgono relazioni diverse da quelle imposte dagli assiomi; essi sono generatori liberi. Con due generatori liberi *a*, *b*, ogni elemento ha un'unica forma ridotta

$$a^{\alpha_0}b^{\beta_1}a^{\alpha_1}b^{\beta_2}a^{\alpha_2}\cdots b^{\beta_k}a^{\alpha_k},$$

dove *k* è un intero non negativo e tutti gli esponenti sono interi non nulli, salvo quelli estremi α_0 e α_k, che possono essere zero. Per $k = 0$ resta una sola potenza di *a*, comprendente l'identità. Un enunciato analogo vale con più generatori.

Scriviamo presentazioni di due gruppi usando insiemi disgiunti di simboli per i generatori. Riunendo generatori e relazioni, senza aggiungere relazioni fra i due insiemi, si ottiene un nuovo gruppo, detto prodotto libero dei due gruppi.

La teoria dei gruppi liberi e, più in generale, dei prodotti liberi occupa un posto importante. Geometricamente, per figure sufficientemente regolari come quelle qui considerate, l'unione di due figure connesse in un solo punto ha come gruppo fondamentale il prodotto libero dei loro gruppi. Il cilindro ha un gruppo libero con un generatore. Ne segue, per esempio, che la figura 24 ha gruppo fondamentale libero con due generatori.

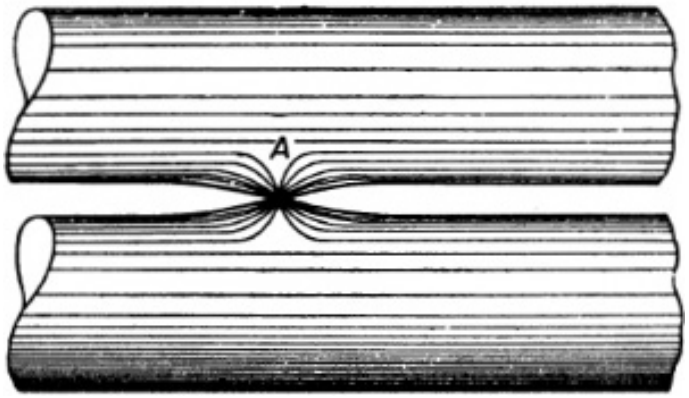


Figura 24.

Come per le superfici, si definisce il gruppo fondamentale anche per corpi spaziali limitati o illimitati.

Nodi e gruppi dei nodi.

Due superfici sono topologicamente equivalenti se esiste fra esse un omeomorfismo. La classificazione delle superfici chiuse era già nota da tempo: ogni superficie chiusa connessa immersa senza autointersezioni nello spazio ordinario è equivalente a una sfera oppure a una sfera con un certo numero di manici (figura 25). Il toro corrisponde a un manico; la superficie del cubo è equivalente alla sfera. Questa classificazione si ottiene anche senza ricorrere ai gruppi fondamentali. Vi sono invece problemi dall'enunciato semplice nei quali essi risultano essenziali; uno dei più celebri è quello dei nodi.

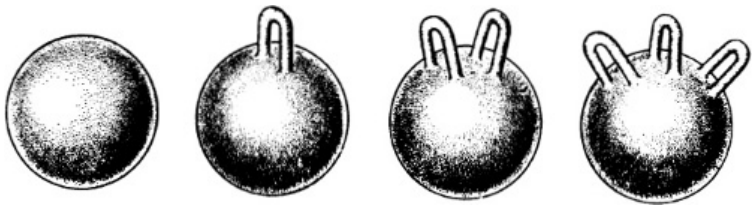


Figura 25.

Un nodo è una curva chiusa semplice nello spazio tridimensionale. La figura 26 mostra quanto possa variare il suo intreccio. Due nodi sono equivalenti se si passa dall'uno all'altro mediante una deformazione continua dello spazio, senza tagliare la curva e senza farla attraversare sé stessa. Sorgono due problemi: riconoscere se due nodi dati mediante proiezioni piane sono equivalenti e classificare i nodi non equivalenti.

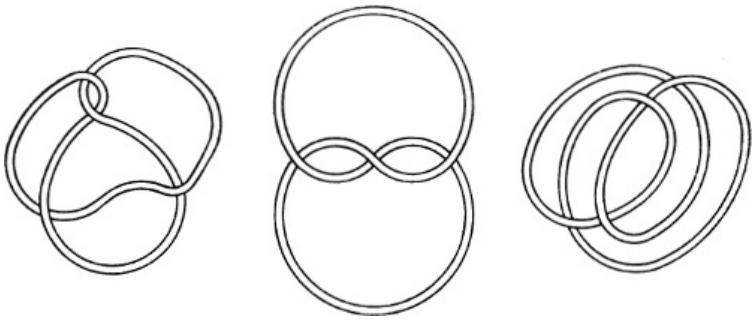


Figura 26.

All'epoca dell'opera questi problemi erano presentati come ancora aperti; importanti risultati parziali erano legati alla teoria dei gruppi.

Togliamo dallo spazio i punti del nodo e consideriamo il gruppo fondamentale dello spazio rimanente: è il gruppo del nodo. Nodi equivalenti hanno gruppi isomorfi; pertanto gruppi non isomorfi provano che i nodi non sono equivalenti. Il nodo equivalente a una circonferenza ha gruppo ciclico infinito, mentre il nodo a trifoglio (figura 27) ha un gruppo non commutativo. I due gruppi non sono isomorfi e il trifoglio non può quindi essere sciolto in una circonferenza senza tagli o attraversamenti. Un fatto intuitivo riceve così una dimostrazione matematica precisa.

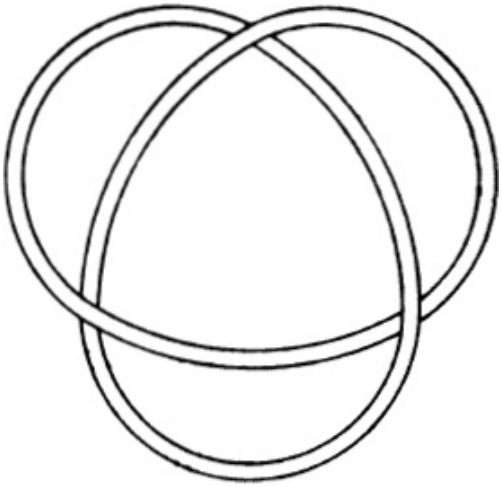


Figura 27.

Anche qui emergono difficoltà profonde. Da una proiezione di un nodo si ricavano facilmente generatori e relazioni del suo gruppo, ma confrontare gruppi dati da presentazioni è un problema difficile. L'autore richiama il risultato di P. S. Novikov sull'impossibilità di un algoritmo generale per decidere determinate questioni di uguaglianza nei gruppi finitamente presentati e ne trae dubbi sulla possibilità di un metodo uniforme per riconoscere l'equivalenza dei nodi. Questa conclusione esprime una valutazione storica: l'indecidibilità per gruppi generali non implica l'indecidibilità per la classe particolare dei gruppi dei nodi, né prova da sola l'impossibilità di un algoritmo per i nodi.

§9. Rappresentazioni e caratteri dei gruppi

La teoria generale dei gruppi somiglia, per il metodo, alla geometria elementare: entrambe partono da un sistema di assiomi. La geometria analitica mostra però quanto possano essere utili i metodi numerici e analitici per affrontare questioni geometriche.

Nella teoria dei gruppi, il collegamento con l'analisi e l'algebra classica è offerto dalla teoria delle rappresentazioni. Come la geometria analitica risolve problemi geometrici con l'analisi e illumina geometricamente problemi analitici, così le rappresentazioni non sono soltanto uno strumento ausiliario per i gruppi. Esse collegano concetti profondi delle due discipline, traducono fatti sui gruppi in relazioni numeriche e danno interpretazioni algebriche a relazioni analitiche. Gran parte delle importanti applicazioni della teoria dei gruppi alla fisica passa proprio attraverso le rappresentazioni.

Rappresentazioni mediante matrici.

Nel capitolo XVI abbiamo studiato la moltiplicazione delle matrici, associativa ma in generale non commutativa. Le matrici quadrate invertibili di ordine fissato formano un gruppo: il prodotto è ancora invertibile, il neutro è la matrice identità e l'inversa di ogni matrice appartiene allo stesso insieme.

Sia G un gruppo. Se a ogni g associamo una matrice complessa invertibile A_g di ordine n in modo che $A_{gh} = A_g A_h$, abbiamo una rappresentazione matriciale di grado n. Si dice di solito semplicemente rappresentazione di grado n: è un omomorfismo di G nel gruppo delle matrici invertibili di quell'ordine. Per le proprietà degli omomorfismi, il neutro va nella matrice identità e gli inversi nelle matrici inverse.

Una matrice di ordine 1 è un numero complesso. Una rappresentazione di grado 1 associa dunque a ogni elemento un complesso non nullo, rispettando il prodotto. L'applicazione che associa 1 alle permutazioni pari e −1 alle dispari è un esempio.

Associando a ogni elemento la matrice identità di ordine n si ottiene la rappresentazione banale di grado n. Un gruppo finito non banale possiede anche rappresentazioni non banali e, facendo variare il grado, se ne ottengono infinite. Vedremo come costruirle.

Da una rappresentazione $g \to A_g$ si possono ottenere rappresentazioni equivalenti cambiando base. Scegliamo una matrice invertibile P dello stesso ordine e poniamo $B_g = P^{-1}A_gP$. La corrispondenza $g \to B_g$ è ancora una rappresentazione, poiché

$$B_{gh} = P^{-1}A_{gh}P = P^{-1}A_gA_hP = P^{-1}A_gPP^{-1}A_hP = B_gB_h\,.$$

Le rappresentazioni ottenute in questo modo si chiamano equivalenti. Non vengono considerate essenzialmente diverse: la teoria le studia normalmente a meno di equivalenza.

Un altro metodo è la somma diretta. Siano $g \to A_g$ e $g \to B_g$ rappresentazioni di gradi m e n. Consideriamo l'applicazione

$$g \to \begin{bmatrix} A_g & 0 \\ 0 & B_g \end{bmatrix}.$$

Per la regola del prodotto delle matrici (capitolo XVI),

$$gh \to \begin{bmatrix} A_{gh} & 0 \\ 0 & B_{gh} \end{bmatrix} = \begin{bmatrix} A_gA_h & 0 \\ 0 & B_gB_h \end{bmatrix} = \begin{bmatrix} A_g & 0 \\ 0 & B_g \end{bmatrix}\begin{bmatrix} A_h & 0 \\ 0 & B_h \end{bmatrix};$$

L'applicazione è quindi ancora una rappresentazione, detta somma diretta delle due precedenti e indicata qui con $A_g + B_g$. Scambiando gli addendi si ottiene

$$g \to \begin{bmatrix} B_g & 0 \\ 0 & A_g \end{bmatrix},$$

che è equivalente alla precedente. A meno di equivalenza, la somma diretta è dunque commutativa e, come si verifica facilmente, associativa. Da rappresentazioni A_g, B_g, C_g, …

possiamo costruirne di grado sempre maggiore: $A_g + B_g + C_g$, oppure quattro copie di A_g, e così via.

Per esempio, $\{1, -1, i, -i\}$ è un gruppo moltiplicativo. Associando ogni numero a sé stesso si ha una rappresentazione di grado 1. Un'altra è data da $1 \to 1$, $-1 \to -1$, $i \to -i$, $-i \to i$. La somma delle due è

$$1 \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, -1 \to \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, i \to \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, -i \to \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}.$$

Coniugandola mediante la matrice

$$P = \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}$$

si ottiene la rappresentazione equivalente

$$1 \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, -1 \to \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, i \to \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, -i \to \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Tutte le sue matrici sono reali, fatto interessante. Supponiamo ora che tutte le matrici di una rappresentazione di grado n abbiano la forma

$$g \to A_g = \begin{bmatrix} B_g & C_g \\ 0 & D_g \end{bmatrix},$$

dove B_g e D_g sono matrici quadrate e il blocco inferiore sinistro è nullo. Moltiplicando A_g e A_h otteniamo

$$A_{gh} = A_g A_h = \begin{bmatrix} B_g B_h & B_g C_h + C_g D_h \\ 0 & D_g D_h \end{bmatrix};$$

dunque $B_{gh} = B_g B_h$ e $D_{gh} = D_g D_h$. Le applicazioni corrispondenti sono rappresentazioni di grado minore, se entrambi i blocchi hanno dimensione positiva. Una rappresentazione equivalente a una di questa forma triangolare a blocchi si dice riducibile. Se non ammette una tale forma con due blocchi non vuoti, si dice irriducibile.

Se anche il blocco superiore destro C_g è nullo per ogni g, la rappresentazione si scinde nella somma diretta di B_g e D_g. Una rappresentazione equivalente a una somma di irriducibili si dice completamente riducibile.

Ogni rappresentazione complessa di grado finito di un gruppo finito è completamente riducibile.[3] Per trovare tutte le rappresentazioni basta dunque conoscere le irriducibili: ogni altra è equivalente a una loro somma diretta.

Il calcolo esplicito delle rappresentazioni irriducibili è in genere piuttosto complesso. L'autore ricorda soluzioni esplicite per particolari classi, fra cui i gruppi commutativi e i gruppi simmetrici, mentre sul piano teorico le proprietà delle rappresentazioni dei gruppi finiti erano già state studiate approfonditamente.

Ogni gruppo finito ha una rappresentazione particolare, detta regolare. Numeriamo gli elementi g_1, ..., g_n in un ordine arbitrario e scriviamo

$$g_i g_k = g_{i_k} \quad (i, k = 1, 2, \cdots, n).$$

Fissato k, costruiamo una matrice di ordine n con un 1 nella posizione di riga i e colonna i_k, per ogni $i = 1, \ldots, n$, e zeri altrove. Indichiamola con R_{g_k}. L'applicazione $g_k \to R_{g_k}$ è la rappresentazione regolare; semplici calcoli verificano che rispetta il prodotto.

Cambiando la numerazione degli elementi si ottiene una rappresentazione equivalente. A meno di equivalenza, ogni gruppo finito ha quindi un'unica rappresentazione regolare.

Enunciamo i teoremi fondamentali per le rappresentazioni complesse dei gruppi finiti. Il numero delle irriducibili non equivalenti è finito ed è uguale al numero delle classi di coniugio (§6). Il grado di ogni irriducibile divide l'ordine del gruppo. La rappresentazione regolare è equivalente alla somma di tutte le irriducibili non equivalenti, ciascuna ripetuta un numero di volte pari al proprio grado.

Ne segue un'interessante relazione fra l'ordine del gruppo e i gradi delle sue rappresentazioni irriducibili.

Siano n l'ordine del gruppo, k il numero delle classi di coniugio e n_1, ..., n_k i gradi delle irriducibili. La rappresentazione regolare ha grado n ed è somma di n_1 copie della prima irriducibile, n_2 della seconda, ecc. Poiché i gradi si sommano nella somma diretta,

$$n = n_1^2 + n_2^2 + \cdots + n_k^2 . \qquad (11)$$

Ogni gruppo possiede la rappresentazione banale irriducibile di grado 1, che associa 1 a ogni elemento. Se nella (11) n_1 è il suo grado, possiamo scrivere

$$n = 1 + n_2^2 + \cdots + n_k^2 ,$$

dove n_2, ..., n_k sono i gradi delle irriducibili non banali.

Sapendo che questi gradi dividono n, talvolta li si determina dalla sola (11), quando k è noto. Il gruppo simmetrico S_3 ha tre classi di coniugio: l'identità; le trasposizioni (12), (13), (23); i cicli (123), (132). Con $n = 6$ e $k = 3$ l'unica possibilità è $6 = 1^2 + 1^2 + 2^2$. Vi sono quindi due irriducibili di grado 1 e una di grado 2.

In un gruppo abeliano finito ogni elemento costituisce una classe di coniugio. Quindi $k = n$ e la (11) impone che tutti i gradi siano 1. Le irriducibili sono tante quanti gli elementi del gruppo.

Le rappresentazioni irriducibili dei gruppi abeliani si chiamano anche caratteri. In generale, il carattere di una rappresentazione è la funzione che associa a ogni elemento la traccia della matrice corrispondente, cioè la somma degli elementi diagonali. I caratteri dei gruppi finiti hanno proprietà e relazioni notevoli. Lo studio di rappresentazioni e caratteri ha arricchito la teoria di risultati generali, con vaste applicazioni alla fisica teorica.

§10. La teoria generale dei gruppi

Per quasi tutto l'Ottocento la teoria si sviluppò soprattutto come teoria dei gruppi di trasformazioni. Gradualmente emerse il ruolo fondamentale dei gruppi astratti e dei loro sottogruppi, ai quali si riconduce lo studio delle azioni per trasformazioni. Il passaggio avvenne dapprima per i gruppi finiti; lo sviluppo dei gruppi di Lie e l'ingresso della teoria in topologia richiesero poi una teoria generale, nella quale i gruppi finiti costituiscono soltanto un caso particolare.

L'autore indica nel libro di O. Ju. Šmidt, pubblicato a Kiev nel 1916, il primo manuale che adottò pienamente questa prospettiva. Negli anni Venti Šmidt ottenne anche un importante teorema sui gruppi infiniti, punto di partenza per ricerche di altri algebristi sovietici. La sua attività e quella di P. S. Aleksandrov, che contribuì alla diffusione dell'algebra moderna, portarono alla formazione di una scuola moscovita poi guidata dal loro allievo A. G. Kuroš. Quest'ultimo divenne noto, fra l'altro, per il teorema secondo cui ogni sottogruppo di un prodotto libero è un prodotto libero di gruppi isomorfi a opportuni sottogruppi dei fattori e, eventualmente, di un gruppo libero aggiuntivo. La sua successiva monografia raccolse sistematicamente i risultati della teoria generale e divenne un riferimento internazionale, che l'autore giudicava il più completo disponibile all'epoca.

Sull'esempio della scuola moscovita, algebristi di Leningrado e di altre città si dedicarono alla teoria generale, contribuendo al suo sviluppo. Le ricerche sovietiche descritte dall'opera coprivano i principali rami della disciplina e avevano ripetutamente esercitato una notevole influenza sulla sua evoluzione.

§11. Numeri ipercomplessi

Risolvendo problemi pratici con metodi algebrici si arriva, nei casi più semplici, a una o più equazioni dalle quali ricavare le incognite. Queste rappresentano caratteristiche quantitative degli oggetti studiati; le equazioni esprimono relazioni reali fra gli oggetti.

Ciò accade per grandezze come massa, volume o distanza, descrivibili da un solo numero. Molti oggetti richiedono però più numeri, talvolta infinitamente molti. Forza, velocità e accelerazione si rappresentano mediante vettori e richiedono tre componenti. La posizione di un punto nello spazio richiede tre coordinate; un piano si descrive localmente con tre parametri, una retta con quattro e un corpo rigido con sei. I problemi relativi a oggetti più complessi conducono quindi a molte incognite, e trattare le equazioni può risultare più laborioso che sfruttare direttamente le proprietà geometriche o fisiche. Nacque così l'idea di descrivere tali oggetti non con sistemi di numeri ordinari, ma con numeri generalizzati sui quali compiere operazioni analoghe a quelle aritmetiche. L'idea era naturale: la storia mostrava la flessibilità del concetto di numero, ampliatosi dai naturali ai razionali, agli algebrici, ai reali e infine ai complessi.

Numeri complessi.

Nel capitolo IV abbiamo studiato le proprietà fondamentali e le prime applicazioni dei complessi. Qui ci interessa fondarne il concetto. Nei reali la radice quadrata di un numero negativo non è definita, poiché ogni quadrato è non negativo. Le esigenze della scienza portarono però a trattare espressioni come $a + b\sqrt{-1}$ come nuovi numeri, chiamati immaginari. Assumendo regole di calcolo coerenti con quelle usuali, tutte le radici quadrate dei numeri negativi si esprimono mediante $i = \sqrt{-1}$, e ogni risultato di un numero finito di operazioni si scrive $a + bi$, con a, b reali.

Una presentazione di questo tipo può apparire contraddittoria: prima si dice che $\sqrt{-1}$, $\sqrt{-2}$, ecc. non hanno significato, poi le si chiama numeri. Ciò indusse molti matematici del Seicento e del Settecento a dubitare della legittimità dei complessi. I dubbi furono superati all'inizio dell'Ottocento grazie alla rappresentazione geometrica mediante punti del piano. Una fondazione puramente aritmetica, dovuta a Bolyai e Hamilton, procede invece nel modo seguente.

Al posto di $a + bi$ consideriamo la coppia di reali (a, b). Due coppie sono uguali se coincidono le componenti: $(a, b) = (c, d)$ se e soltanto se $a = c$ e $b = d$. Definiamo somma e prodotto mediante

$$(a, b) + (c, d) = (a + c, b + d); (a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Per esempio,

$$(2, 3) + (1, -2) = (3, 1), \qquad (2, 3)(1, -2) = (8, -1),$$
$$(3, 0) + (2, 0) = (5, 0), \qquad (3, 0)(2, 0) = (6, 0).$$

Le operazioni sulle coppie con seconda componente zero si riducono alle operazioni sulle prime componenti; possiamo quindi identificare $(a, 0)$ con a. Indicando con i la coppia $(0, 1)$, otteniamo

$$(a, b) = a(1, 0) + b(0, 1) = a + bi,$$
$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1;$$

cioè la notazione usuale dei numeri complessi.

Da questo punto di vista, i complessi sono coppie di reali e le operazioni complesse sono particolari operazioni su tali coppie.

Numeri ipercomplessi.

I successi dei complessi spinsero già nei primi decenni dell'Ottocento a cercare numeri superiori rappresentati da terne, quaterne, ecc. di reali, analogamente ai complessi rappresentati da coppie. Dalla metà del secolo si studiarono molti sistemi particolari, detti ipercomplessi. Fra la fine dell'Ottocento e la prima metà del Novecento se ne sviluppò una teoria generale, con importanti applicazioni in matematica e fisica.

Un numero ipercomplesso di rango n è rappresentato da una n-upla di reali $(a_1, \ldots, a_n)$, che chiameremo coordinate. Due numeri sono uguali se coincidono tutte le coordinate. La somma si definisce naturalmente mediante

$$(a_1 , a_2 , \cdots, a_n) + (b_1 , b_2 , \cdots, b_n) = (a_1 + b_1 , a_2 + b_2 , \cdots, a_n + b_n),$$

in analogia con i numeri complessi.

Si definisce allo stesso modo il prodotto per un reale:

$$a(a_1 , a_2 , \cdots, a_n) = (aa_1 , aa_2 , \cdots, aa_n).$$

Resta da definire il prodotto di due numeri ipercomplessi, con risultato ancora appartenente allo stesso sistema.

Generalizzare il prodotto dei complessi non è immediato: lo si può fare in modi diversi, ottenendo sistemi diversi. Dobbiamo quindi chiarire quali proprietà desideriamo. Vorremmo naturalmente operazioni il più possibile simili a quelle dei reali. Quali sono le proprietà essenziali di queste ultime?

Un esame delle proprietà più usate in algebra conduce all'elenco seguente.

1. La somma di due numeri è univocamente determinata.

2. Il prodotto di due numeri è univocamente determinato.

3. Esiste uno zero tale che $a + 0 = a$ per ogni a.

4. Ogni a ha un opposto x, con $a + x = 0$.

5. L'addizione è commutativa:

$$a + b = b + a.$$

6. L'addizione è associativa:

$$(a + b) + c = a + (b + c).$$

7. La moltiplicazione è commutativa:

$$ab = ba.$$

8. La moltiplicazione è associativa:

$$(ab) \cdot c = a \cdot (bc).$$

9. Vale la proprietà distributiva:

$$a(b + c) = ab + ac, \qquad (b + c)\, a = ba + ca.$$

10. Per ogni a e ogni $b \neq 0$ esiste un unico x tale che $bx = a$.

Queste proprietà furono selezionate attraverso un'analisi accurata, e lo sviluppo dell'algebra ottocentesca ne mostrò l'importanza. Un sistema non banale che le soddisfa si chiama campo. Sono campi i razionali, i reali e i complessi: le operazioni restano nell'insieme e soddisfano le

proprietà 1-10. Esistono infiniti altri campi numerici, ma anche campi di oggetti diversi. Per esempio, le frazioni algebriche, cioè i rapporti di polinomi in determinate variabili, con denominatore non nullo e coefficienti in un campo, si sommano, sottraggono, moltiplicano e dividono rispettando le stesse leggi: formano un campo di funzioni razionali. Si potrebbero dare molti altri esempi. Inizialmente si cercò dunque un prodotto che rendesse campo un sistema ipercomplesso, per estendere ancora il concetto di numero. Si scoprì però che, per estensioni di dimensione finita sui reali con prodotto compatibile con la struttura vettoriale, l'unico campo oltre ai reali stessi è quello dei complessi, di rango 2. Questo risultato attribuisce ai complessi un posto speciale: entro queste ipotesi, non si può andare oltre conservando tutte le proprietà 1-10.

Per costruire sistemi superiori fu quindi necessario rinunciare a una o più proprietà dell'elenco.

Quaternioni.

Il primo sistema ipercomplesso di grande rilievo storico fu quello dei quaternioni, i «numeri quadrupli», introdotto da Hamilton alla metà dell'Ottocento. Soddisfa le proprietà 1-10 salvo la commutatività del prodotto, cioè la 7; nel caso non commutativo valgono entrambe le leggi distributive.

Indichiamo le quaterne (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1) con 1, i, j, k. Poiché

$$(a, b, c, d) = a(1, 0, 0, 0) + b(0, 1, 0, 0) + c(0, 0, 1, 0) + d(0, 0, 0, 1)$$

ogni quaternione si rappresenta univocamente nella forma

$$(a, b, c, d) = a \cdot 1 + b \cdot i + c \cdot j + d \cdot k.$$

Il quaternione 1 sarà l'unità: $1 \cdot \alpha = \alpha \cdot 1 = \alpha$ per ogni α. Poniamo inoltre $i^2 = j^2 = k^2 = -1$ e

$$\begin{aligned} ij &= -ji = k, \\ ik &= -ki = -j, \\ jk &= -kj = i. \end{aligned}$$

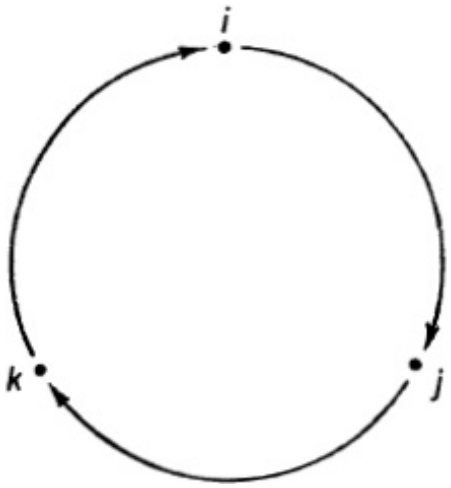


Figura 28.

La figura 28 aiuta a ricordare la tavola: i punti *i*, *j*, *k* sulla circonferenza rappresentano i quaternioni corrispondenti. Il prodotto di due elementi consecutivi nel verso indicato è il terzo; invertendo il verso si cambia il segno. Dalla tavola e dalla distributività si ricava il prodotto di quaternioni arbitrari:

$$
\begin{aligned}
&(a \cdot 1 + b \cdot i + c \cdot j + d \cdot k)(a_1 \cdot 1 + b_1 \cdot i + c_1 \cdot j + d_1 \cdot k) \\
&= aa_1 \cdot 1 + ab_1 \cdot i + ac_1 \cdot j + ad_1 \cdot k \\
&+ ba_1 \cdot i + bb_1 \cdot ii + bc_1 \cdot ij + bd_1 \cdot ik \\
&+ ca_1 \cdot j + cb_1 \cdot ji + cc_1 \cdot jj + cd_1 \cdot jk \\
&+ da_1 \cdot k + db_1 \cdot ki + dc_1 \cdot kj + dd_1 \cdot kk \\
&= (aa_1 - bb_1 - cc_1 - dd_1) \cdot 1 \\
&+ (ab_1 + ba_1 + cd_1 - dc_1) \cdot i \\
&+ (ac_1 + ca_1 - bd_1 + db_1) \cdot j \\
&+ (ad_1 + da_1 + bc_1 - cb_1) \cdot k.
\end{aligned}
$$

Il fattore 1 del termine scalare si omette, scrivendo *a* al posto di $a \cdot 1$. Le uguaglianze $ij = -ji$, $ik = -ki$, $jk = -kj$ mostrano che il prodotto non è commutativo. L'ordine dei fattori va quindi rispettato con cura. Per il resto il calcolo non presenta difficoltà particolari. L'associatività si verifica sulla base 1, *i*, *j*, *k* usando la tavola e si estende al caso generale per distributività.

In $a + bi + cj + dk$, il numero *a* è la parte reale, o scalare, e $bi + cj + dk$ la parte vettoriale. Cambiando segno alla parte vettoriale si ottiene il quaternione coniugato. La somma di due coniugati è reale; moltiplicandoli si trova

$$(a + bi + cj + dk)(a - bi - cj - dk) = a^2 + b^2 + c^2 + d^2; \qquad (12)$$

dunque anche il prodotto è reale.

La somma $a^2 + b^2 + c^2 + d^2$ si chiama qui norma del quaternione. È non negativa e si annulla soltanto per il quaternione nullo. È il quadrato della norma euclidea usuale.

La (12) dice che il prodotto di un quaternione per il suo coniugato è uguale alla sua norma.

Indichiamo il coniugato con un asterisco. Un calcolo diretto verifica

$$(\alpha\beta)^* = \beta^*\alpha^*.$$

Ne segue una proprietà importante: la norma di un prodotto è il prodotto delle norme. Infatti,

$$\begin{aligned}\text{norma}(\alpha\beta) &= (\alpha\beta)(\alpha\beta)^* = \alpha\beta\beta^*\alpha^* \\ &= (\alpha\alpha^*)(\beta\beta^*) = \text{norma}(\alpha)\cdot\text{norma}(\beta).\end{aligned}$$

La norma rende semplice anche la divisione. Per $\alpha = a + bi + cj + dk$ non nullo,

$$\begin{aligned}(a + bi + cj + dk)\frac{1}{a^2+b^2+c^2+d^2}(a - bi - cj - dk) \\ = \frac{1}{a^2+b^2+c^2+d^2}(a^2+b^2+c^2+d^2) = 1;\end{aligned}$$

e quindi il quaternione

$$\frac{1}{a^2+b^2+c^2+d^2}(a - bi - cj - dk) = \alpha^{-1}$$

è l'inverso di α.

Trovato l'inverso, si ricavano i quozienti. Dati $\alpha \neq 0$ e β, dividere β per α significa risolvere una delle equazioni

$$\alpha x = \beta, \qquad y\alpha = \beta.$$

Moltiplicando la prima a sinistra per α^{-1},

$$x = \alpha^{-1}\beta.$$

e moltiplicando la seconda a destra per α^{-1},

$$y = \beta\alpha^{-1}.$$

I prodotti $\alpha^{-1}\beta$ e $\beta\alpha^{-1}$ sono in generale diversi. Bisogna dunque distinguere divisione destra e sinistra; entrambe sono sempre possibili, salvo la divisione per zero.

Algebra dei vettori.

Nonostante molte somiglianze con i complessi, la non commutatività produce differenze profonde. Un'equazione quadratica complessa ha due radici, contate con molteplicità. Ma l'equazione

$$x^2 + 1 = 0$$

nei quaternioni ha già le sei soluzioni $\pm i$, $\pm j$, $\pm k$, e un esame più preciso ne mostra infinite altre. L'autore osserva che queste peculiarità rendono meno immediato l'uso dei quaternioni e giudica il loro ruolo, all'epoca dell'opera, più modesto di quello dei complessi, nonostante i tentativi di Hamilton e di altri di introdurli in matematica e fisica.

I quaternioni diedero però un impulso decisivo all'algebra vettoriale, strumento indispensabile della tecnica e della fisica. Velocità, accelerazione e forza richiedono tre componenti. La parte vettoriale $bi + cj + dk$ di un quaternione è anch'essa determinata da tre numeri e può quindi rappresentare queste grandezze.

Geometricamente essa rappresenta il vettore dall'origine di un sistema cartesiano ortogonale al punto di coordinate b, c, d. Un quaternione si interpreta dunque come coppia formata da uno scalare e da un vettore spaziale. Vediamo come interpretarne le operazioni.

Consideriamo due quaternioni puramente vettoriali $xi + yj + zk$ e $x_1 i + y_1 j + z_1 k$. La somma è ancora vettoriale, con componenti $x + x_1$, $y + y_1$, $z + z_1$. Il vettore risultante è la diagonale del parallelogramma costruito sui due vettori iniziali: ritroviamo la regola del parallelogramma. Analogamente, moltiplicare un quaternione vettoriale per un reale equivale a moltiplicare il vettore per quel reale.

Per il prodotto la situazione è diversa. Infatti,

$$(xi + yj + zk)(x_1 i + y_1 j + z_1 k)$$
$$= -xx_1 - yy_1 - zz_1 + (yz_1 - y_1 z)\, i + (zx_1 - z_1 x)\, j + (xy_1 - x_1 y)\, k;$$

il prodotto di due quaternioni vettoriali è in generale un quaternione completo, con parte scalare e parte vettoriale.

L'opposto della parte scalare del prodotto si chiama prodotto scalare dei vettori; la parte vettoriale dà il prodotto vettoriale. Si indicano rispettivamente con $(\alpha\beta)$, o semplicemente $\alpha\beta$, e $[\alpha\beta]$. Siano i, j, k i vettori unitari degli assi corrispondenti ai quaternioni omonimi. Per $\alpha = xi + yj + zk$ e $\beta = x_1 i + y_1 j + z_1 k$,

$$(\alpha\beta) = xx_1 + yy_1 + zz_1\,,\ [\alpha\beta]$$
$$= (yz_1 - y_1 z)\, \boldsymbol{i} + (zx_1 - z_1 x)\, \boldsymbol{j} + (xy_1 - x_1 y)\, \boldsymbol{k}.$$

Le formule danno anche l'interpretazione geometrica. Il prodotto scalare è il prodotto delle lunghezze per il coseno dell'angolo compreso. Il prodotto vettoriale ha lunghezza uguale all'area del parallelogramma costruito sui vettori ed è perpendicolare al loro piano; il verso segue l'orientazione degli assi: la rotazione dal primo vettore al secondo, vista dall'estremità positiva del prodotto, ha lo stesso verso della rotazione dall'asse x all'asse y vista dall'asse z positivo. Se i vettori sono paralleli, il prodotto vettoriale è nullo.

Nella formulazione usuale della meccanica e della fisica si opera direttamente con i vettori, definendo geometricamente queste operazioni, senza passare ogni volta dai quaternioni.

Concludiamo con un problema meccanico che i quaternioni risolvono con particolare eleganza e che fu fra le motivazioni della loro scoperta.

Un corpo rigido ruota prima di un angolo ϕ attorno a un asse OA e poi di un angolo ϕ_1 attorno a OB, entrambi passanti per O. Quale unica rotazione porta direttamente dalla prima posizione all'ultima? È il problema della composizione delle rotazioni finite. Eulero lo risolse

già nel Settecento con la geometria analitica; i quaternioni ne rendono la soluzione particolarmente trasparente.

Siano $\xi = xi + yj + zk$ variabile e $\alpha = a + bi + cj + dk$ fissato e non nullo. Il quaternione $\alpha^{-1}\xi\alpha$ è ancora puramente vettoriale, come si verifica calcolando. Se ξ, $\alpha^{-1}\xi\alpha$ e la parte vettoriale di α sono rappresentati rispettivamente da $\vec{\xi}$, $\vec{\xi}_1$, $\vec{\alpha}$, allora $\vec{\xi}$ si ottiene da $\vec{\xi}$ mediante una rotazione attorno all'asse di $\vec{\alpha}$. L'ampiezza dell'angolo è descritta dalla formula seguente, scegliendo il verso coerente con la coniugazione adottata:

$$\cos\frac{\phi}{2} = \frac{a}{\sqrt{a^2 + b^2 + c^2 + d^2}}.$$

Il quaternione α rappresenta dunque una rotazione di angolo ϕ attorno all'asse $\vec{\alpha} = bi + cj + dk$. Se la parte vettoriale è nulla, la trasformazione è l'identità.

Viceversa, assegnati asse e angolo, si può trovare un quaternione che rappresenti la rotazione. Ve ne sono infiniti, tutti multipli reali non nulli l'uno dell'altro.

Sia ora una seconda rotazione di angolo ϕ_1 attorno all'asse $\vec{\beta} = b_1 i + c_1 j + d_1 k$., rappresentata da $\beta = a_1 + b_1 i + c_1 j + d_1 k$. La prima rotazione manda $\vec{\xi} = xi + yj + zk$ in $\alpha^{-1}\vec{\xi}\alpha$, e la seconda manda quest'ultimo in $\beta^{-1}(\alpha^{-1}\vec{\xi}\alpha)\,\beta$. Per associatività il risultato si scrive

$$\beta^{-1}(\alpha^{-1}\vec{\xi}\,\alpha)\,\beta = (\alpha\beta)^{-1}\vec{\xi}\,\alpha\beta.$$

La coniugazione del quaternione vettoriale $\vec{\xi}$ mediante $\alpha\beta$ rappresenta ancora una rotazione. La composizione delle rotazioni rappresentate da α e β, applicate in quest'ordine, è quindi rappresentata dal prodotto $\alpha\beta$. Alla composizione delle rotazioni corrisponde la moltiplicazione dei quaternioni.

Oltre alle applicazioni geometriche e fisiche, i quaternioni hanno applicazioni notevoli in teoria dei numeri. Fra i lavori in questo campo ricordiamo quelli di Ju. V. Linnik.

§12. Algebre associative

Definizione generale di algebra.

Abbiamo descritto i numeri ipercomplessi mediante più numeri reali, identificandoli per concretezza con sistemi di reali. Per le indagini teoriche questa prospettiva è troppo ristretta; si è affermata la definizione più generale seguente.

Un sistema S si chiama algebra, o sistema ipercomplesso, su un campo P se sono definite le operazioni seguenti.

a. A ogni a di P e ogni α di S è associato un elemento $a\alpha$ di S, detto prodotto per uno scalare.

b. A ogni coppia α, β di S è associata univocamente la somma $\alpha + \beta$, ancora in S.

c. A ogni coppia α, β è associato univocamente il prodotto $\alpha\beta$, ancora in S.

Queste operazioni devono soddisfare le proprietà seguenti.[4]

1′. $\alpha + \beta = \beta + \alpha$.

2′. $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.

3′. Esiste uno zero θ con la proprietà

$$\alpha + \theta = \alpha,$$

4′. $a(\alpha + \beta) = a\alpha + a\beta$.

5′. $(a + b)\alpha = a\alpha + b\alpha$.

6′. $(ab)\alpha = a(b\alpha)$.

7′. $0 \cdot \alpha = \theta$ e $1 \cdot \alpha = \alpha$, dove 0 e 1 sono lo zero e l'unità del campo P.

8′. Esistono $\alpha_1, \ldots, \alpha_n$ tali che ogni elemento di S si rappresenti univocamente come $a_1\alpha_1 + \ldots + a_n\alpha_n$, con coefficienti in P.

9′. $(a\alpha)\beta = \alpha(a\beta) = a(\alpha\beta)$.

10′. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ e $(\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha$.

Gli elementi del campo arbitrario P prendono il posto dei reali. Per la proprietà 8′, ogni elemento dell'algebra è determinato da n coefficienti di P: possono essere complessi, razionali, reali o elementi di un altro campo.

Le prime otto proprietà significano che S è uno spazio vettoriale di dimensione finita sul campo P (capitolo XVI, §2), detto campo di base dell'algebra. L'opposto di α è $(-1)\alpha$.

Le proprietà 9′ e 10′ si riuniscono nelle formule

$$(a\beta + b\gamma)\,\alpha = a(\beta\alpha) + b(\gamma\alpha),$$
$$\alpha(a\beta + b\gamma) = a(\alpha\beta) + b(\alpha\gamma),$$

che esprimono la linearità del prodotto rispetto a ciascun fattore, cioè la bilinearità.

Il termine «algebra» ha progressivamente sostituito «sistema ipercomplesso», perché gli elementi dei sistemi più generali possono differire molto dai numeri ordinari. Le denominazioni «sistemi ipercomplessi» e «numeri ipercomplessi» sono rimaste soprattutto per esempi classici come i quaternioni.

Le proprietà 1′-10′ non impongono né commutatività né associatività al prodotto, né l'esistenza di un'unità moltiplicativa, né la possibilità di dividere.

Ogni algebra qui considerata possiede una base: elementi $\alpha_1, \ldots, \alpha_n$ rispetto ai quali ogni elemento si scrive univocamente come combinazione lineare a coefficienti in P. Le basi possono essere diverse; hanno tutte lo stesso numero di elementi, detto rango o dimensione dell'algebra. Su un campo infinito, un'algebra non nulla ammette infinite basi.

I complessi, come algebra sui reali, hanno base 1, i. Sono basi anche 2, $3i$, oppure 1, $a + bi$ con $b \neq 0$.

Sia $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ una base di un'algebra su P. Ogni elemento si scrive univocamente

$$\alpha = a_1\epsilon_1 + a_2\epsilon_2 + \cdots + a_n\epsilon_n .$$

Per un altro elemento $\beta = b_1\varepsilon_1 + \ldots + b_n\varepsilon_n$, dalle proprietà 1′-6′ segue

$$\alpha + \beta = (a_1 + b_1)\,\epsilon_1 + (a_2 + b_2)\,\epsilon_2 + \cdots + (a_n + b_n)\,\epsilon_n .$$

Analogamente, per ogni a di P,

$$a\alpha = aa_1\epsilon_1 + aa_2\epsilon_2 + \cdots + aa_n\epsilon_n .$$

Queste formule determinano la somma e il prodotto per scalari. Il prodotto fra elementi dell'algebra deve invece essere specificato. È sufficiente conoscere i prodotti degli elementi della base ε_i: infatti, per le proprietà 9′ e 10′,

$$(a_1\epsilon_1 + a_2\epsilon_2 + \cdots + a_n\epsilon_n)(b_1\epsilon_1 + b_2\epsilon_2 + \cdots + b_n\epsilon_n) = \sum a_i b_j \cdot \epsilon_i\epsilon_j .$$

Ogni prodotto $\varepsilon_i\varepsilon_j$ appartiene all'algebra e si esprime quindi nella base:

$$\epsilon_i\epsilon_j = c_{ij1}\epsilon_1 + c_{ij2}\epsilon_2 + \cdots + c_{ijn}\epsilon_n .$$

I coefficienti c_{ijk} appartengono al campo di base. Il primo indice indica il primo fattore, il secondo il secondo fattore e il terzo l'elemento di base del quale si prende il coefficiente. Si chiamano costanti di struttura: conoscerle significa conoscere tutte le operazioni dell'algebra.

Per un'algebra di rango n vi sono n^3 costanti, una per ogni terna di indici fra 1 e n. Per esempio, i complessi sui reali hanno base 1, i. Dalle uguaglianze

$$1 \cdot 1 = 1 \cdot 1 + 0 \cdot i, \qquad i \cdot 1 = 0 \cdot 1 + 1 \cdot i,$$
$$1 \cdot i = 0 \cdot 1 + 1 \cdot i, \qquad i \cdot i = -1 \cdot 1 + 0 \cdot i,$$

si ricavano le costanti

$$c_{111} = 1,\ c_{112} = 0,\ c_{211} = 0,\ c_{212} = 1,$$
$$c_{121} = 0,\ c_{122} = 1,\ c_{221} = -1,\ c_{222} = 0.$$

Viceversa, assegnati arbitrariamente n^3 elementi c_{ijk} di un campo P, essi definiscono un'algebra di rango n mediante la legge di prodotto $\epsilon_i\epsilon_j = \sum_{k=1}^{n} c_{ijk}\epsilon_k$, estesa per bilinearità.

Le costanti dipendono dalla base scelta. Una stessa algebra può quindi essere descritta da sistemi diversi di costanti di struttura.

Quando considerare due algebre essenzialmente uguali? Due algebre sullo stesso campo P sono isomorfe se esiste una corrispondenza biunivoca che conserva somma, prodotto e moltiplicazione per ciascuno scalare di P. Nella teoria si studiano normalmente a meno di questo isomorfismo.

Si considerano così le proprietà esprimibili mediante le tre operazioni fondamentali, indipendentemente dalla natura concreta degli elementi.

Un isomorfismo manda una base in una base, e le costanti di struttura rispetto alle basi corrispondenti coincidono. Viceversa, due algebre sullo stesso campo con le stesse costanti in basi opportune sono isomorfe.

Hanno un ruolo particolarmente importante le algebre associative, nelle quali $\alpha(\beta\gamma) = (\alpha\beta)\gamma$. Questa sezione ne descrive alcune proprietà. Fra le algebre non associative sono fondamentali quelle di Lie, il cui prodotto soddisfa

$$\alpha\beta = -\beta\alpha, \qquad \alpha(\beta\gamma) + \beta(\gamma\alpha) + \gamma(\alpha\beta) = 0.$$

Esse sono strettamente legate ai gruppi di Lie del §7.

L'algebra delle matrici.

Nei primi tempi dello studio dei sistemi ipercomplessi l'attenzione si concentrò su esempi particolari, alcuni dei quali abbiamo già esaminato. L'algebra delle matrici, fondamentale nella teoria generale, cominciò a essere studiata verso la metà dell'Ottocento. Richiamiamone le operazioni (capitolo XVI, §1).

Una matrice su un campo P è una tabella rettangolare di suoi elementi. Due matrici sono uguali se coincidono gli elementi nelle stesse posizioni. Considereremo soltanto matrici quadrate, con ugual numero di righe e colonne; tale numero è l'ordine.

La somma si esegue elemento per elemento; il prodotto per uno scalare moltiplica ogni elemento per quello scalare. Nel prodotto di due matrici di ordine n, l'elemento di riga i e colonna j è la somma dei prodotti degli elementi della riga i della prima per quelli corrispondenti della colonna j della seconda. Per esempio,

$$\begin{bmatrix} a & b \\ a_1 & b_1 \end{bmatrix} \begin{bmatrix} x & y \\ x_1 & y_1 \end{bmatrix} = \begin{bmatrix} ax + bx_1 & ay + by_1 \\ a_1x + b_1x_1 & a_1y + b_1y_1 \end{bmatrix}.$$

Le ragioni di questa definizione del prodotto sono state spiegate nel capitolo XVI.

Le matrici di ordine n su P si possono dunque sommare, moltiplicare per scalari e moltiplicare fra loro. Calcoli diretti verificano le proprietà 1′-10′ e l'associatività del prodotto. Esse costituiscono quindi un'algebra associativa su P.

L'uguaglianza evidente

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

mostra che le quattro matrici a destra formano una base dell'algebra delle matrici di ordine 2. In generale, indichiamo con ε_{ij} la matrice con 1 nella posizione (i, j) e zeri altrove. Allora

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \sum_{i,j} a_{ij}\epsilon_{ij} ,$$

e le matrici ε_{ij} costituiscono una base. Poiché le matrici ε_{ij} sono n^2, questo è il rango dell'algebra. La tavola di moltiplicazione degli elementi ε_{ij} è

$$\epsilon_{ij} \cdot \epsilon_{jl} = \epsilon_{il} , \qquad \epsilon_{ij} \cdot \epsilon_{kl} = 0, \qquad j \neq k, \qquad i, j, k, l = 1, 2, \cdots, n.$$

L'algebra contiene un'unità: la matrice identità.

Rappresentazioni delle algebre associative.

Un'applicazione fra algebre A e B sullo stesso campo P è un omomorfismo se conserva somma, prodotto e moltiplicazione per scalari. Un omomorfismo di un'algebra associativa A nell'algebra delle matrici di ordine n è una rappresentazione di grado n. Se elementi distinti hanno immagini distinte, la rappresentazione è fedele, cioè un isomorfismo sull'immagine. Le operazioni dell'algebra si traducono allora esattamente in operazioni matriciali. Trovare rappresentazioni è quindi importante. Esamineremo soltanto alcuni procedimenti semplici, ma fondamentali per la teoria generale.

Scegliamo una base $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ di A e un elemento arbitrario α. I prodotti $\varepsilon_1\alpha, \varepsilon_2\alpha, \ldots, \varepsilon_n\alpha$ si esprimono linearmente mediante $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$. Scriviamo

$$\begin{aligned} \epsilon_1\alpha &= a_{11}\epsilon_1 + a_{12}\epsilon_2 + \cdots + a_{1n}\epsilon_n , \\ \epsilon_2\alpha &= a_{21}\epsilon_1 + a_{22}\epsilon_2 + \cdots + a_{2n}\epsilon_n , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \epsilon_n\alpha &= a_{n1}\epsilon_1 + a_{n2}\epsilon_2 + \cdots + a_{nn}\epsilon_n . \end{aligned}$$

A ogni α corrisponde così una matrice $||a_{ij}||$. Un semplice calcolo, usando l'associatività, mostra che si ottiene una rappresentazione, detta regolare. Il suo grado è il rango dell'algebra.

I complessi sono un'algebra reale di rango 2 con base 1, i. Le uguaglianze

$$\begin{aligned} 1 \cdot (a + bi) &= a \cdot 1 + b \cdot i, \\ i \cdot (a + bi) &= -b \cdot 1 + a \cdot i \end{aligned}$$

mostrano che nella rappresentazione regolare $a + bi$ corrisponde alla matrice

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

L'analoga rappresentazione dei quaternioni è

$$a + bi + cj + dk \rightarrow \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}.$$

Queste rappresentazioni sono fedeli. La rappresentazione regolare di un'algebra arbitraria, invece, può non esserlo; è però sempre fedele se l'algebra possiede un'unità.

Ogni algebra associativa può essere inclusa in un'algebra con unità. La rappresentazione regolare di quest'ultima è fedele e, ristretta all'algebra iniziale, resta fedele. Ogni algebra associativa di dimensione finita ammette quindi una rappresentazione matriciale fedele.

Questo procedimento non costruisce tutte le rappresentazioni. Un metodo più fine si basa sugli ideali, concetto fondamentale della matematica moderna.

Un ideale destro è un sottospazio vettoriale I tale che il prodotto di un suo elemento per qualunque elemento dell'algebra appartenga ancora a I. Scambiando l'ordine dei fattori si definisce l'ideale sinistro. Un ideale sia destro sia sinistro è bilatero. Lo zero da solo e l'intera algebra sono ideali bilateri banali. L'esistenza di altri ideali è spesso legata a proprietà interessanti dell'algebra.

Sia I un ideale destro di A con base $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$, in genere più piccola di una base di A. Per ogni α di A, poiché $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ appartengono a I, anche $\varepsilon_1\alpha, \ldots, \varepsilon_m\alpha$ vi appartengono e si esprimono nella base $\varepsilon_1, \ldots, \varepsilon_m$:

$$\begin{aligned} \epsilon_1\alpha &= a_{11}\epsilon_1 + \cdots + a_{1m}\epsilon_m , \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \epsilon_m\alpha &= a_{m1}\epsilon_1 + \cdots + a_{mm}\epsilon_m . \end{aligned}$$

Associando a α la matrice dei coefficienti si ottiene ancora una rappresentazione di A, ora di grado pari alla dimensione di I. Per un ideale proprio il grado è minore di quello della rappresentazione regolare. Gli ideali destri minimi non nulli forniscono gli spazi d'azione più piccoli rispetto all'inclusione, e hanno perciò un ruolo fondamentale nella teoria.

Struttura delle algebre.

Ogni algebra associativa A si rappresenta fedelmente mediante matrici di un certo ordine. Le matrici immagini formano esse stesse un'algebra, generalmente soltanto una parte dell'algebra matriciale completa. Una parte chiusa rispetto alle operazioni è una sottoalgebra. Ogni algebra associativa è dunque isomorfa a una sottoalgebra di matrici.

Il risultato riduce in linea di principio lo studio di tutte le algebre a quello delle sottoalgebre matriciali, ma non ne descrive direttamente la struttura. Una prima risposta generale fu data alla fine dell'Ottocento da F. E. Molin (1861-1941), professore a Dorpat, oggi Tartu, e poi docente a Tomsk attorno al 1900.

Un'algebra non nulla, con prodotto non identicamente nullo, si dice semplice se i suoi soli ideali bilateri sono zero e l'intera algebra. Molin dimostrò che ogni algebra associativa semplice di dimensione finita sui complessi è isomorfa a un'algebra completa di matrici complesse. L'enunciato dell'opera considera esplicitamente il rango almeno 2; il caso di rango 1 con prodotto non nullo è il campo stesso.

Proseguendo queste ricerche, Wedderburn ottenne all'inizio del Novecento risultati che descrivono in modo molto completo la struttura delle algebre su un campo arbitrario.

Un elemento è nilpotente se una sua potenza è zero. Un ideale è nilpotente se esiste un intero s tale che ogni prodotto di s suoi elementi sia zero. Ogni algebra associativa di dimensione finita ha un unico massimo ideale bilatero nilpotente, detto radicale. Se il radicale è zero, l'algebra è semisemplice. Ogni algebra semisemplice è somma diretta di algebre semplici, e il suo studio si riconduce a quello di queste ultime. Un'algebra associativa non nulla si dice algebra di divisione se possiede unità e ogni elemento non nullo è invertibile; equivalentemente, le equazioni $ax = b$ e $ya = b$, per $a \neq 0$, hanno soluzione unica.

Sui complessi la struttura delle algebre semplici è descritta dal teorema di Molin. Su un campo arbitrario P vale il teorema di Wedderburn: ogni algebra associativa semplice di dimensione finita è isomorfa all'algebra completa delle matrici di un opportuno ordine su un'algebra di divisione di dimensione finita su P. La classificazione si riduce dunque alle algebre di divisione. Sui complessi l'unica è il campo complesso stesso, e si ritrova il teorema di Molin.

Sui reali esistono soltanto tre algebre associative di divisione di dimensione finita: i reali, i complessi e i quaternioni. La dimostrazione non è semplice e viene omessa. Per il teorema di Wedderburn, ogni algebra associativa semplice reale di dimensione finita è quindi un'algebra completa di matrici su uno di questi tre sistemi.

I teoremi di Molin e Wedderburn descrivono così le algebre semisemplici. Per quelle con radicale è fondamentale un ulteriore teorema: sotto opportune ipotesi sul campo di base, per esempio su un campo perfetto, un'algebra A con radicale R contiene una sottoalgebra semisemplice L tale che ogni elemento si scriva univocamente $\lambda + \rho$, con λ in L e ρ in R. La sottoalgebra complementare è determinata a meno di un'opportuna coniugazione.

Questi teoremi danno un quadro ordinato dei tipi possibili di algebre associative. Restano da studiare il radicale nilpotente e il modo in cui interagisce con la parte semisemplice. L'autore descrive la teoria delle algebre nilpotenti come ancora in sviluppo.

§13. Algebre di Lie

Accanto alle algebre associative sono state studiate approfonditamente le algebre di Lie, il cui prodotto soddisfa

$$\alpha\beta = -\beta\alpha, \qquad \alpha(\beta\gamma) + \beta(\gamma\alpha) + \gamma(\alpha\beta) = 0.$$

La loro importanza deriva dal legame con i gruppi di Lie (§7), una delle classi fondamentali di gruppi continui, centrali nella geometria moderna. Per l'origine stessa della teoria rivestono particolare interesse le algebre sui reali e sui complessi.

Un esempio semplice è dato dalle matrici quadrate di ordine n, definendo come nuovo prodotto il commutatore $[A, B] = AB - BA$.

Si verifica facilmente che

$$[A, B] = -[B, A],$$
$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

Le matrici quadrate formano quindi un'algebra di Lie rispetto al commutatore. Ogni loro sottospazio vettoriale chiuso rispetto al commutatore è a sua volta una sottoalgebra di Lie.

Rimase a lungo aperta la domanda se ogni algebra di Lie astratta ammettesse una rappresentazione matriciale fedele. Per le algebre di dimensione finita in caratteristica zero la risposta affermativa fu ottenuta nel 1935 da I. D. Ado, allievo di N. G. Čebotarev.

Delineiamo il legame fra gruppi e algebre di Lie, senza entrare nei dettagli, limitandoci al caso matriciale.

Sia L un'algebra di Lie di matrici. A ogni A associamo l'esponenziale $U = e^A = E + A/1! + A^2/2! + \dots$ Nelle vicinanze della matrice nulla e dell'identità, questa applicazione descrive il gruppo di Lie corrispondente. La formulazione è locale: l'insieme di tutti gli esponenziali non è necessariamente un gruppo, e un'algebra non determina da sola un unico gruppo globale. Ogni gruppo di Lie possiede invece un'algebra di Lie, e gruppi con la stessa algebra sono localmente isomorfi. Una formulazione completa richiede i concetti di gruppo locale e isomorfismo locale.

Il passaggio dall'algebra al gruppo avviene dunque localmente tramite un esponenziale, e quello inverso tramite un logaritmo.

Se L è l'algebra di tutte le matrici di ordine n, essa è l'algebra di Lie del gruppo delle matrici invertibili: ogni matrice sufficientemente vicina all'identità si scrive e^A.

Una matrice $A = ||a_{ij}||$ è antisimmetrica se $a_{ji} = -a_{ij}$. Le matrici antisimmetriche reali formano un'algebra di Lie: combinazioni lineari e commutatori restano antisimmetrici.

L'esponenziale di una matrice reale antisimmetrica è ortogonale, e ogni matrice ortogonale sufficientemente vicina all'identità si ottiene così. L'algebra di Lie del gruppo ortogonale è dunque quella delle matrici antisimmetriche.

Ogni rotazione dello spazio attorno all'origine è rappresentata da una matrice ortogonale di ordine 3 con determinante +1; la composizione corrisponde al prodotto delle matrici. Il gruppo delle rotazioni è quindi il gruppo ortogonale speciale di ordine 3. La sua algebra di Lie coincide con quella di tutte le matrici antisimmetriche di ordine 3, cioè delle matrici della forma

$$A = \begin{bmatrix} 0 & -a & -b \\ a & 0 & -c \\ b & c & 0 \end{bmatrix}.$$

Ogni matrice è determinata dai tre numeri a, b, c e si associa al vettore con quelle componenti. Le combinazioni lineari di matrici corrispondono alle combinazioni lineari dei vettori. Il commutatore

$$\begin{aligned} [A_1, A_2] &= A_1A_2 - A_2A_1 \\ &= \begin{bmatrix} 0 & -a_1 & -b_1 \\ a_1 & 0 & -c_1 \\ b_1 & c_1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -a_2 & -b_2 \\ a_2 & 0 & -c_2 \\ b_2 & c_2 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -a_2 & -b_2 \\ a_2 & 0 & -c_2 \\ b_2 & c_2 & 0 \end{bmatrix} \begin{bmatrix} 0 & -a_1 & -b_1 \\ a_1 & 0 & -c_1 \\ b_1 & c_1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & b_2c_1 - b_1c_2 & a_1c_2 - a_2c_1 \\ b_1c_2 - b_2c_1 & 0 & a_2b_1 - a_1b_2 \\ a_2c_1 - a_1c_2 & a_1b_2 - a_2b_1 & 0 \end{bmatrix} \end{aligned}$$

corrisponde al vettore con componenti $b_1c_2 - b_2c_1$, $a_2c_1 - a_1c_2$, $a_1b_2 - a_2b_1$: il prodotto vettoriale dei due vettori. I vettori ordinari, con somma, prodotto per scalari e prodotto vettoriale, formano dunque l'algebra di Lie del gruppo delle rotazioni spaziali attorno a un punto fisso. Si vede così il legame fra i concetti geometrici e le leggi del moto dei corpi rigidi.

Fra la fine dell'Ottocento e l'inizio del Novecento si ottennero per le algebre di Lie risultati analoghi a quelli delle algebre associative, ma con enunciati e prove più complessi. I lavori di Lie, Killing e Cartan portarono ai concetti di radicale e semisemplicità e alla classificazione delle algebre di Lie semplici reali e complesse. Cartan e Weyl svilupparono la teoria delle rappresentazioni matriciali, divenuta uno strumento potente. L'autore ricorda inoltre i risultati ottenuti dai matematici sovietici nei quindici anni precedenti l'opera, sulle rappresentazioni, sulle sottoalgebre semisemplici e sulla struttura delle algebre con radicale assegnato.

§14. Anelli

Nel §11 abbiamo definito un campo mediante le proprietà 1-10. Omettendo la divisione, la commutatività e l'associatività del prodotto si ottiene un anello nel senso ampio, eventualmente non associativo, adottato qui dall'autore; si richiedono entrambe le leggi distributive. Nell'uso moderno, «anello» senza qualificazioni indica di norma un anello associativo.

Ogni campo e ogni algebra, considerando soltanto somma e prodotto, è un anello in questo senso. L'esempio più semplice è l'insieme degli interi con le operazioni usuali. Sono anelli anche i numeri delle forme $a + bi$, $a + b\sqrt{2}$, $a + b\sqrt[3]{2} + c\sqrt[3]{4}$, ecc., con coefficienti interi. Si chiamano anelli numerici; alcune loro proprietà e applicazioni sono state trattate nei capitoli IV e X.

Esistono anelli importanti i cui elementi non sono numeri: i polinomi in variabili assegnate con coefficienti in un anello o campo fissato, le funzioni reali o complesse continue su un dominio e gli endomorfismi di uno spazio vettoriale. Nel caso di uno spazio di Hilbert si può considerare, per esempio, l'anello degli operatori lineari limitati.

Le proprietà aritmetiche degli anelli numerici appartengono alla teoria dei numeri algebrici, al confine fra algebra e teoria dei numeri. Gli anelli di polinomi e i loro ideali sono legati alla geometria algebrica; gli anelli di funzioni e operatori hanno un ruolo fondamentale nell'analisi funzionale (capitolo XIX).

Su queste teorie concrete si svilupparono rapidamente nel Novecento la teoria generale degli anelli e quella degli anelli topologici.

Per ragioni di spazio presenteremo soltanto alcuni risultati introduttivi.

Ideali.

Un sottoinsieme non vuoto I di un anello K, anche non associativo, è un ideale bilatero se contiene la differenza di ogni coppia dei suoi elementi e i prodotti ax, xa per ogni x in I e ogni a in K.

Ogni ideale è a sua volta un anello con le operazioni ereditate. È dunque un sottoanello, senza richiedere che contenga l'eventuale unità dell'anello maggiore. Il contrario non vale in generale.

L'intersezione di ideali è ancora un ideale. In particolare, l'intersezione di tutti quelli contenenti un elemento a è l'ideale principale generato da a, indicato con (a).

Analogamente si definisce l'ideale generato da più elementi. In un anello associativo commutativo con unità, quello generato da $a_1, \ldots, a_n$ è l'insieme delle somme $x_1a_1 + \ldots + x_na_n$, con coefficienti nell'anello. In particolare, (a) è l'insieme dei multipli xa.

Negli interi ogni ideale è principale. Lo stesso vale per i polinomi in una variabile su un campo, per gli interi gaussiani $a + bi$ e per altri anelli. Invece i polinomi razionali in due variabili x, y con termine costante nullo formano un ideale non principale dell'anello di tutti i polinomi razionali in x, y.

Come per i gruppi, si costruisce l'anello quoziente K/I. Due elementi sono congruenti modulo I se $a - b$ appartiene a I; si scrive $a \equiv b \pmod{I}$. La relazione è riflessiva, simmetrica e transitiva, e suddivide K in classi. La somma di due classi è la classe della somma di due rappresentanti; il prodotto è la classe del loro prodotto. Le proprietà dell'ideale garantiscono che il risultato non dipenda dai rappresentanti e che le classi formino un anello.

Gli anelli quoziente svolgono un ruolo analogo ai gruppi quoziente e permettono di costruire nuovi anelli con proprietà diverse. Per esempio, ogni anello commutativo associativo con unità è isomorfo a un quoziente di un anello di polinomi a coefficienti interi, con un insieme di variabili sufficientemente grande.

Proprietà aritmetiche degli anelli.

Nei campi e negli anelli numerici qui considerati, un prodotto è zero soltanto se almeno un fattore è zero. In altri anelli ciò può fallire: due matrici non nulle possono avere prodotto nullo. Se $ab = 0$ con a, b non nulli, essi sono divisori dello zero, rispettivamente a sinistra e a destra. Un anello privo di tali elementi si dice senza divisori dello zero.

Per studiare la divisibilità si considerano di solito anelli associativi commutativi con unità, non nulli e senza divisori dello zero, detti domini d'integrità. Gli anelli numerici citati e gli anelli di polinomi su campi sono esempi.

In un dominio K, b divide a se $a = bq$ con q in K. Somme di multipli di b e prodotti con almeno un fattore multiplo di b sono ancora multipli. Per generalizzare i numeri primi occorre prima distinguere gli elementi associati: due elementi non nulli a, b sono associati se ciascuno divide l'altro. Scrivendo $a = bq_1$ e $b = aq_2$, la cancellazione dà $q_1 q_2 = 1$. I fattori che collegano elementi associati sono dunque unità, cioè elementi invertibili, o divisori dell'unità. Ogni elemento è divisibile per ogni unità. Negli interi le unità sono ± 1; negli interi gaussiani sono ± 1, $\pm i$.

Ogni elemento ammette fattorizzazioni banali $a = (a\varepsilon)\varepsilon^{-1}$, con ε unità. Un elemento non nullo e non invertibile che non ammette fattorizzazioni in due non invertibili si chiama irriducibile. In un dominio generale non va confuso con un elemento primo. È importante determinare le classi di anelli in cui vale una fattorizzazione unica, a meno dell'ordine e di fattori invertibili, analoga a quella degli interi. Essa vale, per esempio, nei domini a ideali principali, cioè nei domini in cui ogni ideale è principale. Si studiano anche analoghi problemi non commutativi.

Il concetto di ideale nacque proprio dallo studio della fattorizzazione. Verso la metà dell'Ottocento, nelle ricerche legate al problema di Fermat, secondo cui $x^n + y^n = z^n$ non ha soluzioni intere non nulle per $n \geqslant 3$, Kummer studiò i numeri $a_0 + a_1\zeta + \ldots + a_{n-1}\zeta^{n-1}$, con coefficienti interi e $\zeta = \cos(2\pi/n) + i \sin(2\pi/n)$. Essi formano un dominio. La fattorizzazione unica, che si sarebbe tentati di trasferire dagli interi ordinari, può però fallire. Per recuperarne

una forma valida, Kummer introdusse fattori «ideali», non necessariamente presenti nell'anello originario. Nella teoria successiva, ai numeri ideali si sostituirono opportuni insiemi di elementi divisibili da essi: gli ideali. La scoperta della non unicità della fattorizzazione negli anelli numerici fu uno dei fatti algebrici più importanti dell'Ottocento e contribuì alla nascita della teoria dei numeri algebrici.

Una notevole applicazione al problema di rappresentare interi come somme di quadrati è stata ricordata alla fine del capitolo X. Contributi importanti alla teoria degli anelli numerici furono dati da E. I. Zolotarev, G. F. Voronoj, I. M. Vinogradov e N. G. Čebotarev.

Varietà algebriche.

Un'altra origine della teoria degli ideali è la geometria algebrica. Studiando le coniche può sorprendere che l'iperbole, pur avendo due rami disconnessi, sia chiamata una sola curva, mentre una coppia di rette sia una conica degenere. L'algebra chiarisce la distinzione: nell'equazione $f(x, y) = 0$, il polinomio dell'iperbole è irriducibile di secondo grado, mentre quello di due rette distinte è prodotto di due fattori lineari. La distinzione fra curve irriducibili e riducibili è dunque algebrica, non semplicemente legata alla connessione dei punti reali.

Per le curve nello spazio la situazione è più complessa. Una curva può essere descritta, nei casi più semplici, da due equazioni $f(x, y, z) = 0$ e $g(x, y, z) = 0$, ma i polinomi non sono determinati univocamente dalla curva; in generale possono occorrere più equazioni. Come definire allora l'irriducibilità?

La risposta naturale viene dagli ideali. Siano $f_1, f_2, \ldots$ polinomi complessi in x, y, z. L'insieme M dei punti dello spazio complesso in cui si annullano tutti si chiama insieme algebrico, o varietà algebrica nel senso qui adottato. I polinomi che si annullano in ogni punto di M formano un ideale I. Esso è radicale: se contiene una potenza di un polinomio, contiene il polinomio stesso. Sistemi diversi di equazioni possono definire lo stesso insieme, ma la corrispondenza fra insiemi algebrici complessi e ideali radicali è biunivoca.

Conviene dunque studiare gli ideali associati, anziché particolari scelte di equazioni. Se I è l'intersezione di due ideali radicali I_1, I_2, il suo insieme algebrico è l'unione dei corrispondenti M_1, M_2. Un insieme algebrico non vuoto è irriducibile se non è unione di due sottoinsiemi algebrici propri; equivalentemente, il suo ideale è primo. La decomposizione in componenti irriducibili corrisponde a un'intersezione di ideali primi. L'esistenza e l'unicità di tali decomposizioni sono fra i primi problemi fondamentali della geometria algebrica e della teoria degli ideali.

Struttura degli anelli non commutativi.

Ogni algebra è anche un anello rispetto a somma e prodotto; molti concetti si estendono quindi agli anelli. Trasferire risultati più fini, come i teoremi di Molin e Wedderburn (§12), richiede però ipotesi e strumenti ulteriori. L'autore descrive i progressi compiuti nei dieci o quindici anni precedenti l'opera. Un problema centrale è definire un radicale che permetta di descrivere gli anelli semisemplici e includa la teoria delle algebre come caso particolare. Diverse nozioni di radicale conducono, sotto opportune condizioni, a teorie di struttura soddisfacenti. L'interesse per gli anelli non commutativi è alimentato anche dagli anelli di operatori dell'analisi funzionale.

§15. Reticoli

Un insieme è parzialmente ordinato se è definita una relazione di precedenza, o subordinazione, con queste proprietà: ogni elemento è subordinato a sé stesso; se a è subordinato a b e viceversa, sono uguali; se a è subordinato a b e b a c, allora a è subordinato a c. La relazione si indica con $\leqslant$.

Un esempio importante è l'insieme di tutti i sottoinsiemi di un insieme dato, ordinati mediante inclusione.

Se ogni coppia è confrontabile, l'ordine è totale, o lineare. È il caso dei reali, dove $a \leqslant b$ significa che a non supera b. L'insieme delle parti di un insieme con almeno due elementi non è invece totalmente ordinato: due sottoinsiemi non vuoti disgiunti non sono confrontabili per inclusione.

Supponiamo che ogni coppia a, b abbia un minimo maggiorante comune c: vale $a \leqslant c, b \leqslant c$ e, per ogni d che soddisfa $a \leqslant d, b \leqslant d$, si ha $c \leqslant d$. L'insieme è allora un semireticolo superiore e c si chiama unione reticolare, indicata qui come «somma» $a + b$. Questa operazione soddisfa

$$a + b = b + a, \qquad (a + b) + c = a + (b + c), \qquad a + a = a. \quad (13)$$

Vale anche il viceversa. Se un'operazione di somma soddisfa le (13), definiamo a subordinato a b quando $a + b = b$. Si ottiene un ordine parziale per il quale $a + b$ è il minimo maggiorante comune.

Analogamente, sostituendo il minimo maggiorante con il massimo minorante comune, si definisce un semireticolo inferiore. Il massimo minorante si chiama intersezione reticolare, indicata qui come «prodotto». Esso soddisfa

$$ab = ba, \qquad (ab)c = a(bc), \qquad aa = a. \qquad (14)$$

Un insieme parzialmente ordinato che sia sia semireticolo superiore sia inferiore è un reticolo. Le due operazioni soddisfano le (13) e (14), ma sono anche legate fra loro: la relazione $a \leqslant b$ equivale sia ad $a + b = b$ sia ad $ab = a$. La compatibilità si esprime con le identità di assorbimento

$$a + ab = a, \qquad a(a + b) = a, \tag{15}$$

Lo studio dei reticoli diventa così lo studio algebrico di due operazioni soggette alle (13), (14), (15). Questo punto di vista permette di esprimere le proprietà di reticoli concreti mediante identità e relazioni algebriche e di utilizzare metodi sviluppati nella teoria dei gruppi e degli anelli.

L'insieme delle parti di un insieme è un reticolo: la somma è l'unione e il prodotto l'intersezione. Anche famiglie particolari di sottoinsiemi possono formare reticoli, con le opportune operazioni di minimo maggiorante e massimo minorante. Sono esempi i sottogruppi di un gruppo, i suoi sottogruppi normali, i sottoanelli di un anello e i suoi ideali. Nei reticoli dei sottogruppi normali e degli ideali vale inoltre la legge modulare:

$$a(ab + c) = ab + ac.$$

I reticoli che soddisfano questa legge, detti modulari o di Dedekind, formano un capitolo importante della teoria.

Molti teoremi sui gruppi e sugli anelli riguardano la disposizione di sottogruppi, sottogruppi normali e ideali. Si possono quindi riformulare come teoremi sui reticoli corrispondenti e, sotto opportune ipotesi, generalizzare. In questo modo risultati importanti sono passati dalle teorie particolari alla teoria dei reticoli. Viceversa, gli strumenti reticolari aiutano a scoprire proprietà dei gruppi e degli anelli concreti.

La teoria dei reticoli si sviluppò soprattutto negli anni Venti e Trenta del Novecento. L'autore la descrive come una disciplina ormai ben formata, ricca di risultati e problemi, pur non avendo ancora applicazioni paragonabili per ampiezza a quelle della teoria dei gruppi.

§16. Altri sistemi algebrici

Abbiamo cercato di mostrare come l'applicazione dei metodi algebrici a problemi sempre più vari abbia ampliato gli oggetti dell'algebra e generalizzato il concetto di operazione. Un ruolo importante spettò allo sviluppo del metodo assiomatico, stimolato fra l'altro dai lavori di N. I. Lobačevskij sui fondamenti della geometria, e alla teoria generale degli insiemi.

Si chiarirono così i concetti generali di operazione e sistema algebrico. Al posto delle operazioni concrete dell'algebra scolastica, prevalentemente numeriche, si considera una regola che a ogni m-upla ordinata $a_1, \ldots, a_m$ di elementi di S associa un unico elemento dello stesso insieme: un'operazione m-aria. Un insieme con una o più operazioni è un sistema algebrico. Studiarli e classificarli è un compito fondamentale, ma nella massima generalità il problema è troppo vasto. Le teorie più feconde riguardano particolari classi. L'autore attribuisce un posto eminente ai gruppi fra i sistemi con un'operazione, e ai campi, alle algebre, agli anelli e ai reticoli fra quelli con più operazioni. Le classi studiate continuano però ad aumentare. Concetti come omomorfismi, sistemi liberi, prodotti liberi, prodotti diretti e radicali vengono estesi a sistemi più generali, dando origine a una nuova branca dell'algebra.

Si sottolinea spesso che nell'algebra la continuità è assente o subordinata, e che la disciplina privilegia quindi il discreto. Questo riflette una sua caratteristica importante. Nel linguaggio filosofico dell'opera, continuo e discontinuo si trovano nel mondo reale in unità dialettica; per conoscere la realtà occorre però talvolta separarne gli aspetti e studiarli distintamente. L'attenzione dell'algebra alle relazioni discrete non va quindi considerata un limite privo di utilità.

La teoria dei gruppi mostra che le discipline algebriche forniscono non soltanto strumenti di calcolo, ma anche un linguaggio per esprimere leggi profonde della natura. Oltre alle applicazioni in fisica, chimica, cristallografia e altre scienze, l'algebra ha un ruolo centrale nella matematica stessa. Richiamando le parole di N. G. Čebotarev, l'autore la descrive come origine di molte idee nuove e come stimolo decisivo per le branche matematiche che costituiscono la base diretta delle scienze fisiche e tecniche.

Letture consigliate

W. Burnside, Theory of groups of finite order, Dover, New York, 1956.

M. Hall, Jr., The theory of groups, Macmillan, New York, 1959.

A. G. Kurosh, The theory of groups, 2 vols., Chelsea, New York, 1960.

D. Montgomery and L. Zippin, Topological transformation groups, Interscience,

New York, 1955. L. S. Pontryagin, Topological groups, Princeton University Press, Princeton,

N. J., 1958. H. Weyl, Symmetry, Princeton University Press, Princeton, N. J., 1952.

Note al capitolo XX

[1] «Finita» significa qui limitata: l'intera figura è contenuta in una regione limitata del piano, per esempio in un disco.

[2] Per i gruppi discreti piani di movimenti propri si veda D. Hilbert e S. E. Cohn-Vossen, *Geometry and the imagination*, Chelsea, New York, 1952. Per la deduzione dei gruppi cristallografici spaziali si veda E. S. Fedorov, *Symmetry of regular systems of figures*, nelle opere raccolte, Accademia delle Scienze dell'URSS, Mosca, 1949.

[3] Ricordiamo che le matrici delle rappresentazioni considerate hanno coefficienti complessi.

[4] Le lettere greche indicano elementi del sistema S; quelle latine elementi del campo P.

Nota editoriale

Sono conservate le sedici sezioni, le 28 figure numerate, le quattro tavole, le formule, la bibliografia e le quattro note. Le tavole numeriche e alcune formule con parole inglesi sono ricomposte in italiano. Si mantiene la convenzione dell'autore: in AB si applica prima A e poi B. Sono corretti refusi negli indici, negli esponenti, nella formula di Viète per il secondo coefficiente, nel teorema di omomorfismo (il quoziente è G/N), nello sviluppo del prodotto quaternionico, nelle costanti di struttura dei complessi e negli assiomi di bilinearità e distributività.

La simmetria di una parte P richiede PA = P, non la sola inclusione PA in P. I rettangoli e i parallelogrammi con il numero minimo di simmetrie si intendono privi di simmetrie aggiuntive. I poliedri regolari considerati sono quelli convessi. Le figure finite del §4 sono limitate; quando non è definito un baricentro si può usare il centro del minimo disco contenitore. Per i laterali si adotta la denominazione corrente: gH è sinistro e Hg è destro. Gli isomorfismi sono biunivoci e gli omomorfismi non sono necessariamente suriettivi, salvo dove si parla esplicitamente di un'applicazione su un gruppo.

Le costruzioni di Galois si svolgono in caratteristica zero; il campo di base contiene i razionali e il campo di spezzamento contiene il campo di base. I cammini sono considerati a meno di omotopia con estremi fissi. Il gruppo fondamentale è invariante per omeomorfismi e, su spazi connessi per archi, non dipende dal punto base a meno di isomorfismo. Il richiamo storico all'indecidibilità per presentazioni di gruppi non implica l'indecidibilità del riconoscimento dei nodi: il problema della parola, quello dell'isomorfismo fra gruppi e quello dell'equivalenza dei nodi vanno distinti.

Le rappresentazioni dei gruppi finiti sono complesse e di dimensione finita. Somma significa somma diretta, non somma ordinaria di matrici. Nella riducibilità entrambi i blocchi hanno dimensione positiva. La rappresentazione regolare usa una matrice di permutazione con un 1 in ogni riga e colonna. Le affermazioni di classificazione sulle algebre associative e di divisione riguardano dimensioni finite sul campo di base. La norma quaternionica dell'autore è il quadrato della norma euclidea; il verso delle rotazioni deve essere coerente con la scelta ξ □ $\alpha^{-1}\xi\alpha$.

Un ideale nilpotente ha nulli tutti i prodotti di un numero sufficientemente grande di suoi elementi: questa proprietà si distingue dalla nilpotenza dei singoli elementi. Per le algebre semplici si esclude il prodotto identicamente nullo. Il teorema di decomposizione con un complemento semisemplice richiede ipotesi sul campo; il complemento è unico soltanto a meno dell'equivalenza appropriata. Il legame fra esponenziale, algebra e gruppo di Lie è locale: gli esponenziali non costituiscono sempre da soli un gruppo e una stessa algebra può corrispondere a gruppi globali diversi. Le rotazioni sono matrici ortogonali di determinante +1.

Il termine anello è usato dall'autore anche senza associatività e senza unità; tali proprietà sono indicate dove necessarie. Per domini, unità e fattorizzazioni si assume un anello associativo commutativo con unità. Elementi primi e irriducibili non coincidono in un dominio arbitrario. In geometria algebrica si distinguono ideali radicali e primi: l'irriducibilità dell'insieme corrisponde alla primalità del suo ideale. Una curva spaziale non è necessariamente definibile da due sole equazioni. Nei reticoli si corregge l'equivalenza a + b = b se e soltanto se ab = a.

Le valutazioni sullo stato delle ricerche, le priorità storiche e i richiami a risultati recenti riflettono il periodo dell'opera. In particolare, i passi sui gruppi finiti semplici, sui nodi, sulle rappresentazioni e sulle applicazioni dei quaternioni non sono presentati come descrizioni dello stato attuale delle discipline.

Indice analitico

Indice ricostruito sul testo italiano: principali argomenti, concetti e risultati. I numeri rimandano alle pagine stampate di questa edizione e sono cliccabili. Le voci sono ordinate alfabeticamente; le sottovoci precisano gli argomenti. L'indice dei nomi segue quello degli argomenti. Non sono riportati i rimandi alla paginazione dell'edizione originale.

A

Abaco, 70

Accelerazione, 59, 81, 84, 125, 159, 212, 234-236, 343, 345, 386, 397, 402, 450-452, 474, 515, 621, 622, 684, 768, 770, 937, 998, 1004

Addizione, 21-23, 27, 29, 38, 39, 47, 67, 70, 136, 177, 192, 237, 293, 301-303, 310, 515, 521, 523, 601-603, 632, 679, 680, 683, 685, 686, 691, 699-701, 705, 710, 716, 737, 738, 978, 981, 1000

Algebra

- assiomatica, 294
- associativa, 1005, 1008-1013, 1021
- astratta, 294
- di divisione, 1011
- di Lie, 988, 1012, 1013
- lineare, 294, 761-763, 768, 770, 789, 805, 861, 921, 924, 941, 944, 951
- matriciale, 1010
- nilpotente, 1011
- semisemplice, 1011
- semplice, 1011, 1021
- teorema fondamentale, 76, 300, 309, 310, 314, 315, 320, 340, 516, 557, 568, 797, 843, 865

Algoritmo, 75, 80, 572, 573, 994

Algoritmo di Euclide, 572

Analisi

- funzionale, 44, 69, 82, 287, 294, 560, 722, 724, 749, 773, 861, 864, 916, 920, 928, 935, 941, 947, 1014, 1017
- matematica, 54, 55, 77, 82, 83, 92, 98-100, 131, 146, 147, 182, 291, 427, 437, 439, 441, 497, 570, 571, 573, 632, 724, 729, 748

Analiticità, 559, 560, 562, 568, 655, 660, 666

Anelli, 117, 155, 382, 596, 891, 892, 1014-1019, 1021

Anello

- commutativo, 1015
- dei polinomi, 1015
- quoziente, 1015

Angolo

- di parallelismo, 819, 821-823, 885
- fra vettori, 777, 778, 780, 921, 926

Antinomie, 760

Applicazione

- affine, 253, 257-264, 270, 271, 279, 281, 283, 489, 539, 541, 547, 568, 776, 841, 943
- biunivoca, 21, 273, 274, 733, 734, 736, 737, 739, 760, 770, 846, 896, 915, 979, 982, 986, 990, 1007
- continua, 863, 911-913, 915, 916
- identica, 959, 962, 964, 967
- inversa, 231, 285, 957, 958, 962, 983
- lineare, 293, 385, 792-795, 800, 802, 803, 861, 872, 943, 951
- omotopa, 913, 1021
- proiettiva, 272, 274-277, 283-285, 489, 837-839
- topologica, 891, 892, 894, 901

Approssimazione

- delle funzioni, 68, 69, 140, 632, 633, 649, 671, 920, 942
- in media quadratica, 633, 648, 661-664, 931
- migliore, 577, 633, 647-649, 651-655, 660-663, 721
- polinomiale, 641, 651
- uniforme, 633, 647-649, 651, 652, 756

Arco

- lunghezza, 98, 442-444, 448, 450, 452, 470, 479, 491, 729, 865, 869, 941

Aree, 31, 43, 52, 80, 81, 108, 141, 142, 144, 154, 164, 199, 203, 256, 347, 398, 456, 464, 466, 470, 492, 496, 497, 514, 616, 618, 723, 755, 855

Aritmetica, 16, 19-23, 26-33, 35-38, 40, 43-45, 47, 48, 50, 54, 65, 67, 74, 75, 80, 83, 86, 147, 292, 424, 463, 570, 571, 574, 579, 589, 590, 592, 612, 613, 694, 695, 702, 704-708, 712, 713, 732, 962, 968, 969, 999

Aritmetizzazione, 197, 235

Armoniche, 179, 314, 357, 397-400, 410, 417-419, 432, 510, 655, 656, 660, 669, 686, 937-939, 949

Ascissa, 87, 104, 105, 107, 120, 133, 162, 192, 197, 206, 208, 218, 219, 234, 241, 246, 333-336, 356, 444, 505, 618, 642, 649, 655, 754, 755, 835, 898, 899, 925

Asintoti, 283, 533, 826

Assioma

- della scelta, 760
- delle parallele, 18, 812, 833, 834
- di Archimede, 742
- di continuità, 316, 834

Assiomi

- della probabilità, 619
- sistema, 842, 994

Associatività, 738, 962, 981, 1002, 1005, 1006, 1008, 1009, 1014, 1021

Autofunzioni, 383, 411, 412, 414, 415, 920, 940, 941, 944, 946, 947

Automorfismo, 886, 958, 979, 987

Autovalori, 796-798, 803, 806, 810, 920, 940, 941, 944-947, 950

Autovettori, 796-798, 803, 804, 810, 944, 945

B

Baricentro, 164, 208, 209, 216, 257, 1021

Base

- di uno spazio vettoriale, 773, 774, 794, 796
- di un'algebra, 1006, 1007, 1009
- ortonormale, 779, 780, 791, 794, 802-804, 810, 940, 950

BESM, 692, 693

Biettività, 21, 273, 274, 285, 541, 542, 733, 734, 736, 737, 739, 760, 770, 827, 831, 832, 841, 842, 846, 865, 886, 891, 896, 915, 957, 958, 962, 964, 979, 982-984, 986, 987, 990, 1007, 1016

Bilinearità, 1006, 1007, 1021

Binomio di Newton, 49, 50, 522

Brachistocrona, 494, 495, 504, 514

C

Calcolatrici

- analogiche, 685, 686
- digitali, 685, 686
- elettroniche, 692

Calcolo

- approssimato, 49, 128, 309, 332, 339, 369, 546, 633, 642, 929, 947
- delle variazioni, 60, 82, 430, 432, 487, 494, 497-500, 502, 505, 508, 511, 514, 871, 913, 914, 916, 920, 941, 942
- differenziale, 43, 44, 57-59, 80, 82, 83, 95, 141, 143-146, 196, 207, 292, 299, 343, 394, 437, 438, 453, 491, 728, 761, 817, 841, 920, 928, 947
- integrale, 36, 46, 47, 58, 61, 68, 80, 81, 141, 144, 145, 176, 196, 292, 299, 344
- numerico, 108, 441, 668, 679, 721, 789
- vettoriale, 237

Campi

- estensione, 857
- teoria, 82, 261, 910, 911

Campo

- dei numeri complessi, 516, 517, 520, 521, 523, 555, 565, 567, 810, 1011
- dei numeri razionali, 306, 978
- dei numeri reali, 515, 516, 526, 555
- ordinato, 739, 742
- vettoriale, 386, 395, 396, 567, 908, 910-912

Carattere

- di un gruppo, 913, 994, 997

Caratteristica

- di Eulero, 901, 903, 907, 911

Cardinalità, 734, 736, 737, 743, 748, 749, 752, 760

Carica elettrica, 40, 146, 434

Catena

- di Markov, 723

Causa ed effetto, 623, 631, 882

Centro

- di gravità, 164, 207-209, 216, 257, 970, 1021
- di simmetria, 956

Cerchio, 31, 32, 56, 81, 85, 98, 104, 199, 205-207, 210, 212, 225, 226, 247, 255-257, 276, 284, 285, 307, 308, 315-317, 320, 328, 473, 491, 544, 545, 568, 729, 806, 827-832, 835-838, 842, 845, 862, 863, 877, 879, 880, 886, 974

Ciclo

- in topologia, 905-907
- limite, 381-383, 385, 908, 909

Cicloide, 61, 504, 514

Cilindro, 132, 243, 245, 246, 250, 252, 436, 454, 455, 459, 474, 483, 484, 486-488, 534-537, 892, 900, 901, 930, 953, 989-992

Circonferenza, 32, 55, 56, 98, 104, 142, 164, 205, 206, 210, 255, 258, 285, 313, 328, 346, 364, 373, 376, 381, 436, 444, 447, 450, 469, 471, 472, 474, 476, 483, 491, 518-520, 534-536, 545, 547, 548, 552, 553, 555, 556, 558, 562, 568, 632, 823-825, 827, 829, 830, 832, 838-840, 842, 843, 862, 863, 868, 873, 879, 886, 891, 897-900, 908, 910, 911, 919, 972, 973, 989, 993, 1002

Circuiti

aritmetici, 715

Classe

di equivalenza, 990

di funzioni, 99, 139, 151, 169, 344, 347, 431, 432, 560, 634, 653, 665, 682, 729, 753, 757, 758, 810, 861

Classificazione

affine, 260, 261

dei gruppi, 988

delle superfici, 260, 992

Codice, 695, 697, 698, 700-703, 705-707, 709, 711, 712, 714-717, 727

Coefficiente

di correlazione, 629, 631

di Fourier, 658, 659, 662, 663, 686

Combinazione

lineare, 512, 755, 772-775, 790, 791, 905, 931, 1006, 1012, 1013

Combinazioni, 28, 50, 113, 136, 279, 332, 336, 347, 355, 359, 410, 434, 512, 603, 654, 656, 682, 683, 693, 694, 714, 715, 719, 755, 772-775, 790, 791, 857, 886, 904, 905, 931, 1006, 1012, 1013

Commutatività, 738, 969, 1001, 1003, 1006, 1014

Compattezza, 939

Complemento algebrico, 784, 786

Completezza, 432, 513, 514, 743, 874, 889, 933, 934, 940, 947, 950

Computer, 725

Condizione

al contorno, 398, 399, 401, 403, 408-410, 413, 414, 416, 420, 430, 432, 498, 503, 504, 506, 512, 669, 943, 945

iniziale, 352, 353, 355, 356, 361, 365, 367-369, 377, 378, 382, 383, 401, 403, 405, 408, 409, 412, 413, 428-430, 507, 619, 623, 667, 677, 678, 696, 865, 910

Conduzione del calore, 63, 82, 314, 391, 398, 400, 414, 416, 425, 567, 634, 667, 867, 874

Congruenza, 33, 485, 846, 880

Coniche, 46, 47, 56, 57, 75, 196, 225, 268, 276, 436, 484, 799, 897, 1016

Connessione

riferimenti generali, 22, 26, 45, 64, 68, 143, 144, 175, 196, 293, 568, 827, 863, 874, 889, 894, 896, 906, 919, 1016

semplice, 380, 543, 544, 551, 552, 560, 568

Cono, 36, 225, 245, 246, 248-250, 253, 281-284, 290, 436, 485, 486, 953

Conservazione

dell'energia, 391, 396, 403-405, 936, 940, 945

Continuità

di una funzione, 99, 100, 666

Contrazione, 212, 253-259, 262, 277, 279-281, 283, 690, 691, 841

Convergenza

assoluta, 180, 181, 522, 583

di una serie, 178, 179, 188, 194, 518, 665, 674, 730, 806, 931, 949

in media, 931

quasi ovunque, 753, 760

uniforme, 181-185, 187, 189, 653, 657, 658, 757

Convessità, 125-127, 193, 438, 488, 649, 889

Coordinate

cartesiane, 87, 215, 228, 238, 257, 276, 281, 286, 288, 310, 440, 443, 453, 478, 829, 840, 845, 868, 887, 896

curvilinee, 467, 478, 481

omogenee, 896

polari, 310, 375, 381, 415, 542

sferiche, 493

trasformazione, 231, 792, 795

Corda vibrante, 393, 395, 402-404, 406, 413, 427, 634, 655-657, 661, 665, 930, 946

Corpo

convesso, 487, 832, 861, 892

rigido, 262, 360, 880, 998, 1004, 1013

solido, 82, 456, 534

Correlazione, 629, 631

Corrispondenza

biunivoca, 21, 273, 274, 733, 734, 736, 737, 739, 760, 770, 846, 896, 915, 979, 982, 986, 990, 1007

Coseno, 112, 238, 241, 312, 382, 389, 414, 417, 480, 482, 523, 777, 778, 921, 922, 926, 927, 929, 946, 950, 1004

Costruzioni

con riga e compasso, 305, 307, 308, 571

Cristallografia, 33, 34, 67, 487, 574, 799, 800, 887, 951, 976, 978, 1019

Curva

algebrica, 383

chiusa, 169, 326, 327, 382, 527, 543, 564, 861, 891, 892, 894, 897, 900, 910, 918, 919, 993

di livello, 242

integrale, 362, 364, 366-377, 380-383, 503, 508, 908-910

piana, 104, 157, 438, 440, 441, 443-445, 447, 448, 450, 452, 483, 491, 495, 505, 842

spazio, 243, 438, 440, 441, 444, 445, 451, 452, 484, 491, 505, 508, 870, 1016, 1021

Curvatura

riferimenti generali, 81, 207, 437-439, 441, 446-453, 457-463, 465-469, 471-477, 482, 486, 491-493, 761, 856, 872, 873, 882, 883, 885, 889

gaussiana, 463, 466, 467, 472, 473, 476, 477, 482, 492, 872, 873, 885

geodetica, 471, 474

media, 463, 465, 466, 493

principali, 459, 462, 463, 466, 467, 477

raggio, 81, 207

D

Deduzione, 17, 28, 32, 620, 724, 816, 818, 823, 834, 977, 1020

Definizione

assiomatica, 778

Densità

di probabilità, 625, 626, 630

Derivata

di una funzione composta, 151

parziale, 156-161, 163, 164, 166, 344, 378, 380, 386, 388, 391, 395, 397, 417, 418, 423, 433, 479, 492, 507, 510, 514, 524, 622, 625, 722-724, 761, 792, 799, 871, 889, 916, 920, 935

seconda, 125-128, 193, 334, 399, 401, 421, 422, 427-432, 451, 461, 489, 514, 555, 621, 644, 645, 800

Derivate, 80, 81, 83, 104, 106-110, 112-131, 133, 134, 136-138, 144-152, 156-161, 163, 164, 166, 181, 186, 187, 193, 194, 200, 203, 220, 321-325, 332, 334, 340-342, 344, 351-354, 356-358, 361, 366, 368, 369, 371, 378, 380, 385, 386, 388, 391, 393, 395, 397, 399, 401, 408, 411, 417, 418, 421-423, 427-433, 435, 438, 443, 451, 453, 461, 478, 479, 485, 489, 491, 492, 497, 499, 501-503, 506, 507, 509, 510, 514, 523-526, 528, 530, 532, 535, 539, 541, 542, 547, 548, 551, 554-556, 561, 562, 620-622, 625, 639-641, 644, 645, 652, 654, 657, 660, 666, 673, 675, 678, 696, 720, 722-724, 761, 792, 799, 800, 807, 871, 874, 889, 916, 920, 931, 935, 943

Determinante, 264, 276, 278, 279, 281, 290, 293, 331, 332, 371, 372, 421, 432, 541, 679, 680, 691, 781-789, 792-794, 800, 810, 841, 935, 1013, 1021

Diagramma, 121, 127, 156, 336, 340, 849, 889

Differenza

di insiemi, 732

Differenze finite, 133, 419, 421, 424, 432, 435, 670, 722, 723

Differenziabilità, 987

Differenziale

riferimenti generali, 129-133, 147, 160, 194, 479, 480, 499, 761, 887

totale, 761

Dilatazione, 88, 246, 540, 542, 546, 672, 841, 882

Dimensione

di uno spazio, 773, 863, 900

infinita, 14, 18, 287, 773, 811, 861, 924-926, 942-944, 949

Discontinuità, 102, 103, 109, 183, 372, 428, 665, 752, 861

Diseguaglianza

di Čebyšëv, 612, 613

di Schwarz, 777, 810

triangolare, 864, 915

Disposizioni, 86, 123, 292, 372, 377, 839, 857, 860, 874, 908, 962, 975, 1018

Distanza, 31, 33, 52, 55-59, 77, 78, 80, 81, 84, 85, 87, 88, 96, 99, 104, 107, 108, 110, 120, 131, 132, 142-144, 146, 154, 163, 164, 168, 194, 197, 199, 203, 205, 210, 212, 214, 218, 223, 234, 238, 246, 253, 258, 259, 264, 269, 271, 272, 280, 283, 314, 315, 328, 356, 373, 387, 393, 408, 417, 418, 441, 443, 446, 449, 453, 456, 462, 463, 468-470, 518, 529, 534, 535, 556, 604, 606, 619, 621, 623, 624, 665, 672-674, 686, 691, 743, 744, 776,

820-824, 826, 840, 846, 847, 850, 851, 855, 857, 858, 860-862, 864, 865, 867-875, 878, 879, 882, 883, 885-887, 914-916, 919, 921, 925, 926, 931, 936, 942, 943, 946, 950, 955, 958, 964, 968, 969, 974, 998

Distributività, 239, 492, 732, 738, 1002, 1021

Distribuzione

binomiale, 604

di probabilità, 611, 613, 618

uniforme, 727

Divisibilità, 571, 594-596, 1015

Divisore

dello zero, 1015

primo, 575, 592

Dominio

di definizione, 87, 91, 161, 170, 194, 380, 405, 486, 497, 942

Dualità, 913, 918

E

Elemento

neutro, 981

Ellisse, 46, 56, 57, 81, 196, 199, 204, 209-212, 216, 217, 223-226, 228, 229, 255-257, 260, 266, 268, 269, 276, 286, 287, 373, 436, 547, 548, 568, 839, 841, 891

Ellissoide, 244, 246, 252, 253, 892, 924, 944

Energia

cinetica, 360, 403, 404, 406, 429, 507, 799, 875

potenziale, 404, 429, 432, 435, 496-498, 510, 514, 661

Equazione

algebrica, 47, 196-199, 204, 207, 208, 288, 291-293, 295, 299, 302, 304, 305, 307-309, 355, 421, 424, 432, 515, 516, 567, 670, 767, 978, 987

caratteristica, 354-356, 376

cubica, 49, 296-299, 303, 304

del calore, 400, 402, 414, 425, 426

delle onde, 402, 404, 405, 411, 420, 428

di continuità, 390, 391, 395, 397

di Eulero, 502-504, 508, 514

di Laplace, 397-399, 401, 405, 415, 416, 423, 424, 428, 434, 538, 670, 686-688

di Poisson, 397, 398, 408

di quarto grado, 204-206, 298, 299, 304

di quinto grado, 300, 302, 307

di Schrödinger, 431

di secondo grado, 204, 228, 231, 233, 244, 252, 265, 295, 304

integrale, 68, 69, 82, 388, 419, 424, 560, 567, 670, 672, 675-677, 722, 723, 866, 915, 920, 929, 935, 938-944, 950

lineare, 353, 355, 357, 359, 762, 768, 781, 787, 830, 840, 847, 861

trascendente, 336

Equazione differenziale

alle derivate parziali, 344, 386, 388, 391, 395, 397, 423, 433, 510, 622, 625, 722, 723, 916, 920, 935

lineare, 353, 770, 798, 807, 808, 941

ordinaria, 342, 344, 351, 353, 355, 369, 377, 383, 386, 416, 622, 672, 703, 726, 805, 920, 941

sistemi, 351, 369, 371, 383, 506, 620, 798, 807, 910

soluzione generale, 81, 352-357, 359, 377, 406, 416, 503, 514, 638, 790, 980

Equazioni

di Cauchy-Riemann, 525, 528, 532, 538, 541, 546, 548

Errore

di arrotondamento, 678

relativo, 160, 727, 845, 869

Esistenza

di soluzioni, 81, 365, 667, 790

e unicità, 365, 722

Esponenziale, 353, 355, 385, 520, 521, 631, 806, 824, 889, 1012, 1013, 1021

Estremale, 487

Estremo

inferiore, 744, 747, 750, 755, 871

superiore, 149, 182, 194, 741, 742, 744, 746, 747, 760, 865

Eulero

formula dei poliedri, 522, 901, 907

Eventi

incompatibili, 601

indipendenti, 616

Evento

certo, 601, 602, 618

impossibile, 618

F

Fascio

di rette, 898

Fattorizzazione, 300, 301, 322, 595, 596, 1015, 1016, 1021

Flusso, 77, 82, 173, 174, 176, 286, 389, 391, 396, 398, 400, 434, 528, 529, 531-537, 544, 545, 568, 624, 719

Forma

canonica, 228, 229, 231-233, 244, 260, 266, 267, 802, 804, 808

differenziale, 391, 393-395

fondamentale, 395, 455, 482

quadratica, 244, 293, 480-482, 574, 799, 800, 802-804, 869, 870, 872, 887

quadratica definita positiva, 802, 870-872, 883

Formula

del binomio, 522, 579

di Cauchy, 549, 554

di interpolazione, 633, 681

di Simpson, 642-645

di Taylor, 125, 136-140, 163, 164, 193, 371, 449, 463, 481, 568, 632, 633, 638, 639, 673

di Viète, 300, 301, 978, 1021

Frazione

continua, 721

decimale, 39, 49, 50, 570, 632

razionale, 153, 570, 580

Frequenza

di un evento, 607

Frontiera, 382, 745, 863, 895, 972

Funzionale

riferimenti generali, 14, 44, 69, 70, 72, 82, 88, 287, 294, 342, 343, 350, 431, 432, 497-499, 505, 511, 514, 560, 669, 670, 686, 722, 724, 749, 773, 792, 861, 864, 866, 887, 916, 920, 926, 928, 931, 935, 941, 942, 947, 1014, 1017

continuo, 942

Funzione

algebrica, 566, 569, 861, 912

analitica, 139, 188, 189, 517, 523, 532, 538, 539, 542-544, 546, 548, 549, 551, 552, 555-557, 560-562, 565, 567, 570, 639, 653, 654, 805, 808, 918, 947

approssimante, 632, 648

armonica, 397-400, 418, 419, 432, 510, 656

caratteristica, 532, 535, 536, 544, 545, 755-757, 760

complessa, 44, 60, 68, 82, 285, 435, 515, 516, 523-525, 528, 529, 532, 535, 537-539, 544, 549, 551, 554-556, 562, 567, 573, 578, 832, 886, 912

composta, 91, 151, 461, 478, 543, 682, 691

continua, 55, 69, 99, 100, 103, 110, 123, 148, 169, 181, 184, 185, 309, 427, 430, 431, 633, 634, 641, 649, 651, 653, 654, 657, 660, 662-664, 728, 729, 748, 753, 760, 770, 779, 842, 861, 862, 864-866, 908, 915, 928, 934, 942

crescente, 582

derivabile, 110, 133, 150, 181, 499, 514, 539, 554, 761

di Dirichlet, 753, 755, 760

di Eulero, 589

di Green, 935-937, 949

di variabile reale, 64, 68, 69, 103, 169, 321, 517, 618, 619, 635, 728-730, 920, 928, 929

differenziabile, 943

discontinua, 100, 102, 103, 148, 184, 428, 665, 729, 753, 861

elementare, 100, 113, 116, 118, 119, 151, 152, 190, 347, 359, 517, 526, 545, 558, 653, 702, 720, 806

ellittica, 567

esponenziale, 353, 355, 520, 521, 806

implicita, 160, 342

integrabile, 169, 660, 664, 669, 941

intera, 556-560, 562, 568, 578

inversa, 115, 116, 344, 385, 682, 810

limitata, 148

lineare, 88, 194, 195, 532, 761-763, 784, 792, 810

misurabile, 753-755, 757, 758, 760, 928

ortogonali, 634, 864, 930

periodica, 352, 660, 758

primitiva, 27, 145, 151

razionale, 525, 557-559, 979, 1001

semplice, 118, 353, 359, 635

sferica, 930, 940

spazio, 669, 862, 865, 928

trigonometrica, 88, 113, 139, 347, 520, 526, 658, 663, 681

univoca, 362, 564

zeta, 558

G

Genere

di una superficie, 566

Geodesiche, 471, 473-475, 477, 486, 492, 826, 834, 861, 868, 871, 873-875, 879, 880, 889, 913

Geometria

affine, 840, 841, 857, 883

algebrica, 287, 288, 1014, 1016, 1021

analitica, 47, 54-57, 60, 80, 82, 83, 196-199, 201, 207-209, 228, 234, 235, 241, 253, 260, 261, 269, 286-288, 292, 293, 299, 323, 390, 437, 440, 453, 492, 768, 776, 799, 802, 846, 847, 850, 857, 896, 899, 921, 994, 1005

di Lobačevskij, 816, 818-821, 823-827, 829, 831-834, 838, 839, 842, 856, 873, 883, 885, 886, 889

di Riemann, 867, 868, 872-876, 889

differenziale, 59, 287, 437-441, 448, 455, 458, 476, 483, 485, 486, 488, 489, 491, 856

euclidea, 66, 71, 816, 818, 820, 821, 824, 825, 832-834, 838, 839, 841, 847, 856, 868, 872, 873, 877, 879, 880, 886

intrinseca, 468-473, 475, 477, 478, 480-482, 488, 492, 826, 827, 868, 869, 872, 874, 885

non euclidea, 18, 66, 285, 811, 815, 817, 818, 826, 833, 838, 858, 876, 885, 896, 912, 918

proiettiva, 46, 60, 270, 275, 285, 489, 839, 840, 896, 897, 987

sferica, 46

Gradiente, 396, 434, 514, 722

Grado

di un polinomio, 139

di un'equazione, 231, 233, 288

Grafico

di una funzione, 87, 88, 90, 91, 104, 109, 120, 124-128, 133, 141, 195, 370, 640, 653

Gravità, 84, 125, 144, 159, 207, 208, 346, 494, 621-623, 627, 970

Gruppo

abeliano, 970, 984, 997

alterno, 966, 967, 970

ciclico, 968, 969, 982, 993

commutativo, 738, 913, 996

continuo, 913, 952, 987, 988, 1012

cristallografico, 970, 976, 977, 1020

delle permutazioni, 309, 965, 966, 978-980, 982

delle rotazioni, 965, 1013

di Fedorov, 34, 970, 977

di Galois, 306-309, 340, 978-980

di Lie, 294, 987, 988, 998, 1008, 1012, 1021

di simmetria, 964-966, 970, 973-975, 984

di trasformazioni, 266, 489, 865, 867, 882, 886, 961, 964, 965, 981-983, 989, 998

discreto, 952, 974, 977, 1020

finito, 952, 964, 965, 980, 982, 984, 988, 995-998, 1021

fondamentale, 989-993, 1021

libero, 992, 998

ordine, 964, 965, 984, 997

ortogonale, 1013

quoziente, 983-986, 1015

rappresentazione, 947, 996, 1021

risolubile, 984

semplice, 984

simmetrico, 965, 967, 970, 979, 980, 983-985, 996, 997

teoria, 67, 293, 294, 577, 912, 951, 952, 975-978, 984, 987, 988, 992-994, 998, 1018, 1019

topologico, 913, 988

I

Ideale

riferimenti generali, 1010, 1011, 1014-1016, 1018, 1021

bilatero, 1010, 1011, 1014

nilpotente, 1021

primo, 1016

principale, 1014, 1015

Idrodinamica, 60, 99, 528, 538, 567, 624

Indipendenza

lineare, 432, 434, 669, 772

stocastica, 617, 629

Indirizzo

nelle calcolatrici, 697-702, 706-709, 714, 720, 727

Induzione

completa, 324

Inerzia
momento, 214-217
Infinitesimi, 43, 54, 59, 75, 78, 92-94, 98, 130-132, 159, 176, 178, 470, 630, 761, 762, 858, 867-869
Informazione, 599, 708
Insieme
aperto, 746, 747, 749
chiuso, 746-748, 912-915
di Cantor, 749
finito, 732-735, 746, 750, 955, 966
infinito, 64, 343, 351, 564, 732-735, 746, 760, 964
limitato, 309, 744, 950
misurabile, 751, 754
numerabile, 734, 747, 748, 759
perfetto, 748, 751
vuoto, 730, 731, 746, 760, 863
Insiemi
teoria, 64, 68, 71, 729, 730, 745, 746, 751, 760, 914
Integrale
curvilineo, 173, 175, 509, 549
definito, 142-145, 148, 169, 444, 549, 642, 643, 702, 941
di Cauchy, 552, 554
di Fourier, 946
di Lebesgue, 169, 729, 754, 755, 757, 758, 760, 928, 934, 950
di Riemann, 754, 755
doppio, 172, 175, 456
ellittico, 152
indefinito, 149-151, 153, 551
multiplo, 169, 172, 174, 508, 721
primo, 910
volume, 390
Integrazione
numerica, 633, 720, 723
per parti, 152, 153, 429, 501, 583
Interpolazione
di Lagrange, 639
Intersezione
di insiemi, 731
Intervalli, 43, 58, 78, 86, 88, 90, 92, 96, 99, 100, 104, 115, 120-126, 133-139, 141-146, 148, 149, 155, 169, 176, 181-189, 194, 302, 316, 322, 323, 332, 377, 388, 389, 392, 401, 414, 421, 422, 426, 428, 495, 499-503, 505, 506, 508, 511, 512, 527, 568, 579, 585, 588, 589, 592, 593, 597, 616, 617, 622, 624, 625, 627, 633, 635, 637-655, 657, 659, 661, 665, 666, 669, 670, 673-676, 691, 693, 696, 697, 699, 704, 705, 715, 723, 728, 736, 740, 742-755, 758, 760, 864, 870, 882, 883, 891, 914-916, 919, 925, 928, 930, 931, 937, 942, 949
Intorno, 17, 49, 75, 163, 186, 188, 316-318, 366, 371, 372, 375-377, 381, 383, 385, 392, 441, 453, 455, 465, 484, 486, 491, 495, 496, 501, 517, 523, 527, 528, 534-537, 539, 541, 544, 545, 552, 555, 557, 561, 563-565, 615, 616, 619, 638, 639, 654, 666, 794, 805, 866, 868, 871, 887, 910
Invariante, 265-267, 269-271, 293, 475, 489, 565, 691, 853, 867, 903, 944, 958, 959, 965, 975, 983, 987, 991, 1021
Inversione, 81, 385, 682, 720, 723, 796, 987, 988
Inviluppo, 483-485, 853
Iperbola, 46, 56, 88, 204, 209, 210, 218, 223-226, 228, 229, 248, 260, 268, 269, 276, 283, 286, 841, 849, 897, 1016
Iperboloide, 244-246, 248-253, 260, 261, 290, 475
Iperpiano, 847-849, 853, 861, 886, 889
Ipotesi
di Riemann, 578
Isometria, 958
Isomorfismo, 774, 982, 984, 986, 990, 1007-1009, 1012, 1021
Iterazione, 424, 672, 675, 677, 691, 722

L

Laplaciano, 399, 400, 428, 434
Legge
dei grandi numeri, 607, 613, 627, 631
del moto, 52, 59, 77, 99, 125, 144, 145, 212, 343, 404, 874, 883, 1013
di gravitazione, 81, 212, 287, 617
Limite
di una successione, 192, 753, 805
inferiore, 143, 162, 332
superiore, 143, 672
Logaritmo
riferimenti generali, 49, 50, 55, 70, 75, 97, 112, 113, 292, 526, 527, 564, 565, 568, 570, 582, 584, 655, 681, 690, 721, 885, 886, 982, 1012
naturale, 112, 526, 690, 885

Logica
matematica, 70, 71, 724, 725, 749

Lune, 31, 77

Lunghezza, 31, 35, 37-40, 70, 78, 84, 88, 98, 123, 141, 142, 159, 169, 176, 212, 213, 215, 219, 235-239, 241, 254, 256, 257, 259, 286, 310, 315-317, 319, 328, 346, 347, 366, 386, 394, 404, 413, 429, 436, 441-444, 446-450, 452, 453, 464, 468-473, 475, 479, 480, 491, 511, 549, 550, 618, 632, 650, 654, 656, 687, 696, 699, 724, 729, 739, 742, 746, 749-752, 755, 776, 778-780, 794, 803, 820, 824-829, 847, 850, 856, 858, 865, 867-871, 873, 875, 879, 880, 883, 885, 887, 889, 910, 921-924, 926, 928-930, 934-937, 941, 944, 946, 949, 966, 967, 1004

M

Macchine
a programma, 684, 693, 695, 696, 698-703, 705, 720-722, 724, 725, 727
memoria, 681, 684, 694, 695, 697, 698, 700, 702, 706-709, 712-714, 716-719, 722
unità aritmetica, 694, 702, 704-708, 712, 713
unità di comando, 694, 695, 698, 706, 708

Massimo
comune divisore, 341, 572, 573
e minimo, 59, 61, 81, 120, 121, 126, 160, 164, 332, 333, 405, 462, 498, 657, 799, 942

Matematica
applicazioni, 342
astrazione, 14, 15, 18, 27-31, 33, 34, 37, 38, 40, 41, 49, 54, 66, 67, 71, 74, 82, 99, 427, 729, 811, 841, 850, 865, 874, 876, 878, 882, 883, 890, 898
storia, 45, 48, 50, 65

Matrici
diagonale, 798
inversa, 788, 995
moltiplicazione, 765-767, 785, 995, 1013
ortogonale, 802, 804, 1013, 1021
quadrata, 763, 766, 767, 783, 785, 788, 793, 805, 806, 995, 996, 1008, 1012
rango, 790
simmetrica, 799, 800

Meccanica
celeste, 377, 379, 809, 900, 909
classica, 920
quantistica, 18, 69, 83, 292, 294, 431, 864, 920, 941, 945, 947, 949

Media
aritmetica, 424, 463, 612, 613
quadratica, 633, 648, 661-665, 931, 933, 934
valori medi e speranza matematica, 106, 133, 134, 138, 193, 340, 611-613, 626, 627, 723

Metodo
delle corde, 334
delle differenze finite, 421, 424, 670
delle successive approssimazioni, 365, 367, 424, 672, 673, 675, 677
di Eulero, 365, 511, 670, 678, 696-698
di Fourier, 409, 416
di Newton, 334-336
di Ritz, 432, 669
di separazione delle variabili, 409, 414, 420
variazionale, 722

Misura
di Lebesgue, 750
di un insieme, 749-751, 757
nulla, 760

Modello
di Beltrami, 867
di Klein, 831, 833, 835
di Poincaré, 886

Modulo
di un numero complesso, 310, 592
di un vettore, 479, 492

Moto
browniano, 616, 626
periodico, 88, 365, 382, 656, 900
uniforme, 451, 875

N

Nastro di Möbius, 895-897, 902

Nodo, 274, 375, 380, 414, 424, 637, 639-641, 645, 646, 651, 655, 666, 670, 688, 817, 908, 992-994, 1021

Norma
euclidea, 1002, 1021

Numerazione
binaria, 703-705

decimale, 23, 26, 49, 703, 705

Numeri

algebrico, 293, 573, 597, 598, 910, 911, 913, 1014, 1016

complesso, 14, 300, 310-312, 314, 315, 317-319, 326, 327, 340, 515, 516, 526, 532, 542, 565, 568, 592, 595, 776, 780, 865, 978, 995, 999, 1000

composto, 575

di Betti, 906, 907, 913

intero, 14, 20, 21, 27, 35, 36, 38, 42, 44, 47, 50, 54, 60, 92, 111, 192, 292, 306, 307, 334, 414, 570-572, 574, 597, 732

ipercomplesso, 294, 951, 998-1000, 1005, 1006, 1008

irrazionale, 37, 47, 192, 632

naturale, 93, 96, 143, 177, 570, 572, 575, 581, 589, 594, 650, 730, 732-736, 742, 955, 957

negativo, 43, 47, 48, 247, 310, 526, 999

primo, 27, 47, 74, 305, 308, 558, 570-572, 574-579, 587, 594, 1015

razionale, 35, 37, 148, 152, 192, 294, 300, 306, 632, 653, 735-739, 978

reale, 37-43, 52, 63, 68, 75, 86, 97, 98, 154, 155, 229, 237, 300, 310, 316, 322, 325, 497, 515, 516, 526, 570, 622, 653, 728, 737, 739-741, 743, 776, 778, 803, 921, 927, 928, 1005

teoria, 26, 27, 38, 41, 43, 44, 47, 51, 60, 71, 97, 293, 310, 314, 487, 558, 568, 570-575, 580, 593, 594, 596, 632, 635, 724, 728, 776, 799, 800, 845, 853, 860, 1005, 1014, 1016

Numeri primi

distribuzione, 578

gemelli, 575

infinità, 579

O

Omeomorfismo, 912, 919, 990-992, 1021

Omologia, 905, 913

Omomorfismo, 984-986, 995, 1009, 1019, 1021

Omotetia, 794

Omotopia, 913, 1021

Onda, 17, 62, 402, 404-408, 411, 414, 420, 425, 428, 434, 666, 717, 727, 879, 940, 941, 950

Operatore

autoaggiunto, 941, 945-947

compatto, 940, 944, 950

integrale, 939, 940, 943, 945

limitato, 945

lineare, 567, 940, 941, 943, 944, 1014

Operazione

algebrica, 44, 48, 113, 572, 722, 753

Ordinata, 87, 162, 192, 197, 202, 208, 234, 241, 414, 505, 511, 512, 571, 587, 598, 754, 755, 835, 981, 1019

Ordine

di un gruppo, 964, 965, 984, 997

di un'equazione differenziale, 352, 429

Orientabilità, 874, 901-903, 906, 907, 912, 919

Ortogonalità, 265, 411, 412, 434, 658, 662, 666, 669, 779, 922, 923, 944

Oscillazioni

armoniche, 357, 410, 655, 656, 937, 938

forzate, 358, 359, 939

libere, 358, 937-940

proprie, 356, 359, 411-414, 427, 930

Osculatore

piano, 441, 451-453, 461, 491, 493, 856

P

Parabola, 46, 56, 57, 75, 81, 92, 131, 143, 196, 199, 204-206, 209, 210, 219-226, 228, 229, 249, 250, 260, 268, 269, 276, 286, 290, 321, 332, 350, 367, 436, 635, 638, 642, 839, 841

Paraboloide, 131, 132, 221, 242, 243, 245, 249, 250, 253, 453

Paradosso, 37

Parallelismo, 489, 704, 776, 819, 821-823, 885

Parallelogramma, 67, 173, 235, 257, 261, 264, 480, 481, 515, 649, 768, 841, 959, 1004, 1021

Parametro, 269, 287, 290, 308, 353, 361, 362, 370, 382, 383, 409, 410, 414, 420, 441, 478, 483-485, 500, 501, 620-624, 647-649, 661, 668-671, 685, 708, 720, 725, 824, 825, 856, 861, 870, 886, 899, 940, 949, 987, 988, 998

Pendolo, 35, 159, 346, 347, 351, 352, 360, 369, 377, 385, 898, 899

Periodo, 416, 523, 588, 656, 657, 659, 663-665, 758

Permutazioni

riferimenti generali, 301-303, 306-309, 906, 918, 957, 958, 960, 961, 963, 965-967, 970, 978-980, 982-985, 995, 1021

dispari, 966, 967
pari, 906, 918, 966, 970, 985, 995

Piano
complesso, 312, 313, 315, 316, 318, 321, 323, 326, 516, 518, 520, 543, 843
di fase, 370, 371, 376, 382
euclideo, 827
iperbolico, 823, 826, 827, 885, 912
proiettivo, 273-275, 284, 285, 837, 886, 896, 897, 902, 903, 905, 906, 919
tangente, 282, 455-458, 460-463, 469-471, 474, 475, 480-482, 484, 868, 869, 885

Poliedri, 436, 485, 487, 488, 850-853, 860, 863, 887, 901, 903, 965, 1021

Poliedro
convesso, 487, 860, 887, 901, 903
regolare, 850, 853, 965, 1021

Poligono, 81, 98, 193, 236, 305, 307, 308, 313, 423, 424, 470, 472, 571, 604, 688, 851-853, 891, 904, 905, 919, 953, 973

Polinomi
caratteristico, 797, 798
di Čebyšëv, 640, 646, 651, 652
di Taylor, 666, 691
radici, 318, 321-323, 325-328, 331, 334, 337, 338, 342, 557, 797
trigonometrico, 634

Polo
di una funzione, 518, 558, 559, 562, 568, 578, 598

Posizione generale, 272, 274

Postulato
delle parallele, 812, 833

Potenza
di un insieme, 734, 736, 737, 743, 748, 749, 752, 760

Potenziale, 348, 349, 395, 396, 404, 406, 417-419, 429, 432, 434, 435, 496-498, 507, 510, 514, 531, 532, 535-537, 568, 661

Primitiva, 27, 145, 148, 149, 151, 194, 551, 943

Principio
di Hamilton, 429, 431, 435, 514
variazionale, 507

Probabilità
condizionata, 602, 616
teoria, 61, 71, 152, 599, 601, 606-608, 613, 619, 629

Problema
ai limiti, 400, 402, 420, 431, 503, 506, 510, 669
di Dirichlet, 432
di Neumann, 434

Processo
casuale, 620, 625, 723

Prodotto
diretto, 1019
scalare, 238, 239, 241, 265, 435, 480, 492, 777-780, 794, 803, 810, 922, 923, 927-929, 932, 945, 1004
vettoriale, 1004, 1013

Programma
di Erlangen, 841, 876, 882

Proiezione
ortogonale, 779, 780, 794

Prolungamento
analitico, 561, 562, 578

Proporzioni, 857

Pseudosfera, 826, 827, 868, 885

Punto
aderente, 862, 863, 914-916
all'infinito, 543, 568, 886, 896-898
angoloso, 429
di accumulazione, 97, 745, 746, 910
di flesso, 127, 438
di frontiera, 745, 972
fisso, 111, 244, 393, 441, 447, 452, 673, 866, 874, 911-913, 915, 916, 918, 919, 974, 1013
interno, 284, 371, 398, 399, 402, 499, 519, 635, 652, 666, 746, 799, 827, 829, 837, 845, 863
isolato, 745, 746, 748
regolare, 517
sella, 374
singolare, 371, 372, 375-377, 380-383, 385, 486, 559, 908, 910, 911

Q

Quadratura, 80, 81, 339, 359, 645, 646, 670

Quadriche, 799

Quasi ovunque, 669, 753, 758, 760, 928, 934, 942, 950

Quaternioni, 1001-1006, 1010, 1011, 1021

R

Radice
dell'unità, 302, 306
Radici
multiple, 301, 322, 325, 341, 354, 798
risoluzione per radicali, 965, 978, 980
Raggio
di convergenza, 188, 189, 520
Rango
di una matrice, 790
Rapporto
anarmonico, 840, 886
Rappresentazione
conforme, 285, 541-546, 568, 886
Regione
semplicemente connessa, 380, 544, 551, 552, 560
Regolarizzazione, 566
Relatività, 18, 66, 811, 816, 817, 832, 867, 876, 877, 879-883, 888
Residuo, 137, 139, 193, 598, 623, 624, 684
Reticolo, 261, 419, 421, 423-426, 435, 439, 478, 487, 574, 1017-1019, 1021
Retta
proiettiva, 906
tangente, 107, 125, 133, 455
Riflessione, 30, 220, 222, 264, 265, 282, 302, 456, 476, 794, 831, 836, 841, 886, 956-959, 964, 965, 970-975, 977
Risolvente, 302, 308
Risonanza, 358, 359
Rotazione, 52, 131, 213-215, 221, 231, 232, 244, 252, 267, 282, 441, 461, 495, 500, 514, 531, 540, 542, 617, 685-687, 717, 794, 797, 810, 812, 819, 827, 831, 834, 836, 874, 875, 947, 951, 953, 955-959, 961-963, 965, 968-970, 972-974, 985, 1004, 1005, 1013, 1021

S

Scalare, 237-239, 241, 265, 386, 387, 395, 396, 428, 435, 478, 480, 492, 770, 771, 773, 774, 776-780, 794, 803, 810, 874, 922, 923, 927-929, 932, 945, 1002, 1004, 1005, 1007-1009, 1013
Seno, 46, 49, 112, 116, 117, 312, 412, 414, 523, 946, 950
Serie
a termini positivi, 180
armonica, 179
convergente, 180, 181, 411, 556, 583
di Fourier, 634, 654, 655, 659, 660, 664, 666, 758, 864
di potenze, 186, 188-190, 194, 516-520, 523, 554-557, 560, 568, 570, 654, 668, 720, 805, 806
di Taylor, 139, 140, 186-188, 475, 517, 555, 558, 560, 561
divergente, 180, 185, 194
geometrica, 519, 520, 556, 666, 674, 806
sommabilità, 573
trigonometrica, 359, 567, 634, 656-658, 730, 758, 932, 949, 950
Sezione
conica, 46, 47, 56, 57, 75, 196, 225, 276, 436
di Dedekind, 740
normale, 458-465, 482, 491, 493
Sfera, 27, 32, 33, 46, 77, 154, 155, 246, 287, 383, 436, 454, 460, 465, 467-469, 473, 478, 483, 485, 486, 488, 489, 491-493, 566, 569, 803, 827, 832, 842, 847, 849, 850, 855, 856, 860, 868, 873, 879, 886, 887, 892, 894, 896, 899-901, 905, 913, 919, 930, 953, 977, 989, 990, 992
Similitudine, 199
Simmetria
asse, 220, 221, 286, 952, 953
assiale, 956
centrale, 953, 956, 959
piana, 952
piano, 952
speculare, 952
Simplesso, 851-853, 886, 903, 906
Sistema
algebrico, 294, 670, 676, 679, 688, 693, 951, 989, 1019
completo, 602, 903
di equazioni lineari, 723, 762, 768, 781, 789, 790, 806, 848, 861
dinamico, 383, 622, 900, 910, 913
ortogonale, 662, 664, 929-931, 933, 940
ortonormale, 663, 930-934, 949

Somma
diretta, 995-997, 1011, 1021
Sottogruppo
riferimenti generali, 967, 968, 970, 982-986, 988, 998, 1018
normale, 983-986, 988, 1018
Sottospazio, 772, 775, 779, 780, 790, 791, 794, 796, 942, 1010, 1012
Spazio
affine, 842, 857
delle fasi, 371, 380, 625, 630, 858-860, 887, 900, 910
di dimensione infinita, 924-926, 942, 944, 949
di Hilbert, 864, 924-926, 928-934, 940, 942, 943, 945, 947, 949, 950, 988, 1014
euclideo, 66, 776, 778, 779, 791, 794, 802, 811, 844, 864, 867-869, 872, 873, 885, 887, 912-914
funzionale, 69, 72, 514, 866, 926, 931
lineare, 770, 771, 773, 775, 778, 780
metrico, 864, 865, 915, 916
proiettivo, 66
topologico, 842, 844, 862, 914, 916
vettoriale, 769, 770, 1006, 1014
Spettro
riferimenti generali, 887, 938, 946, 947, 950
continuo, 946, 950
Spirale, 381, 908
Stabilità, 55, 68, 192, 307, 343, 377-380, 385, 613, 615, 616, 619, 671, 677, 693, 703, 723, 839, 841, 941, 986
Successione
convergente, 753
di Cauchy, 741
di funzioni, 169, 182, 511, 512, 632, 653, 748
Superficie
chiusa, 465, 485, 486, 493, 892, 895-897, 901-903, 906, 907, 911, 912, 919, 992
di Riemann, 564-566, 569, 844, 912
di rotazione, 495, 500
minima, 161, 466, 493
orientabile, 919
regolare, 473, 476, 486-488, 869, 885, 889
sviluppabile, 484

T

Tangente
trigonometrica, 106, 107, 157, 438
Tensore, 43, 294, 387, 762, 872, 873
Teorema
centrale del limite, 614, 631
del punto fisso, 919
della media, 421, 422, 444, 673
di Abel, 307, 518
di Bolzano-Weierstrass, 317, 318, 746
di Brouwer, 912
di Cauchy, 551, 552, 568
di Lagrange, 984
di Pitagora, 36, 40, 55, 199, 210, 255, 443, 444, 868, 885, 922, 923, 928, 933, 934
di Rolle, 322, 324
di Weierstrass, 634, 652-654
fondamentale dell'algebra, 76, 300, 309, 310, 314, 315, 320, 340, 516, 557, 568, 797, 843, 865
Teoria
degli insiemi, 64, 68, 71, 729, 730, 745, 746, 751, 760, 914
dei gruppi, 67, 293, 294, 577, 912, 951, 952, 975-978, 984, 987, 988, 992-994, 998, 1018, 1019
dei numeri, 26, 27, 38, 41, 43, 44, 47, 51, 60, 71, 97, 293, 310, 314, 487, 558, 568, 570-575, 580, 593, 594, 596, 632, 635, 724, 728, 776, 799, 800, 845, 853, 860, 1005, 1014, 1016
della misura, 618, 619, 629, 754
della probabilità, 61, 71, 152, 599, 601, 606-608, 613, 619, 629
delle funzioni, 17, 44, 60, 64, 68, 69, 82, 103, 169, 228, 285, 287, 314, 419, 528, 529, 538, 539, 543, 551, 554, 557, 560, 562, 566, 567, 573, 578, 580, 618, 632, 634, 635, 728-730, 749, 752, 753, 808, 832, 886, 912, 918, 920, 928, 929, 947
delle rappresentazioni, 947, 952, 994, 1013
Topologia
combinatoria, 905, 913
insiemistica, 912-915
Toro, 383, 842, 844, 892, 894, 895, 898, 899, 905, 911, 919, 989, 990, 992
Torsione, 441, 452, 453, 474, 491-493, 905

Traiettoria, 52, 343, 369, 372, 379, 436, 474, 529, 604, 616, 617, 623, 647, 692, 703, 856, 910, 919

Trasformazione

affine, 253, 257-264, 270, 271, 279, 281, 283, 489, 539, 541, 547, 568, 776, 841, 943

conforme, 285, 541-546, 886

inversa, 231, 285, 957, 958, 962, 983

lineare, 293, 385, 792-795, 800, 802, 803, 861, 872, 943, 951

ortogonale, 244, 264-266, 794, 800, 802, 804, 810

proiettiva, 272, 274-277, 283-285, 489, 837-839

topologica, 891, 892, 894, 901

Traslazione, 202, 230-232, 244, 267, 542, 770, 828, 829, 834, 837, 851, 882, 887, 955-959, 961-964, 968-970, 974, 985

Trasporto

parallelo, 493

Trasposizione, 49, 768, 966, 967, 997

Triangolazione, 901-907, 918, 919

Triangolo

sferico, 473

U

Unicità, 272, 285, 365, 366, 385, 397, 401-403, 405, 428, 431, 434, 561, 562, 596, 674, 722, 727, 789, 791, 837, 838, 919, 1016

Uniformità, 29, 61, 84, 846

Unione

di insiemi, 863, 914, 916

Unità immaginaria, 515

V

Valore

assoluto, 92, 164, 181, 237, 313, 318, 332, 338, 378, 447, 448, 461, 466, 514, 597, 605, 615, 639, 648-650, 652, 654, 665, 666, 672, 674, 691, 705, 706, 777, 804, 829, 926, 949, 985

atteso, 611-613, 626, 627, 723

medio, 106, 133, 134, 138, 193, 340, 627

Variabile

casuale, 611, 612, 614, 617, 629

complessa, 17, 44, 60, 68, 82, 140, 285, 314, 515, 516, 520, 523, 551, 567, 573, 578, 597, 832, 886, 912

dipendente, 85, 99, 116

indipendente, 85, 86, 99, 106, 107, 116, 130, 132, 133, 160, 321, 342, 351, 365, 369, 377, 386, 387, 413, 420, 423, 432, 552, 629, 683, 761, 807, 887, 941, 955, 979

reale, 64, 68, 69, 103, 169, 314, 321, 517, 525, 539, 618, 619, 635, 728-730, 920, 928, 929

Varianza, 600, 604, 611-615, 623, 624

Variazione

prima, 430, 431, 435

Varietà

algebrica, 293, 1016

differenziabile, 987

topologica, 900, 919

Velocità, 17, 52, 53, 57-59, 62, 70, 78, 80, 81, 83, 88, 95, 102, 104, 106-108, 111, 125, 142-146, 158, 167, 172-174, 176, 234-236, 277, 278, 280, 294, 342-346, 352, 357, 360, 386, 388-390, 392-397, 402, 406, 407, 428, 431, 441, 445, 450-452, 494, 495, 515, 528-537, 545, 568, 620-623, 626, 630, 631, 655, 684, 685, 692, 702-704, 719, 725, 727, 768, 770, 799, 859, 874, 875, 879, 883, 887, 900, 951, 998, 1004

Vettori

applicato, 770

nullo, 775, 933

somma, 769, 922, 925, 962

tangente, 478, 480, 869, 910

unitario, 803, 908

Vibrazioni, 62, 63, 396, 414, 716

Volume, 31, 33, 36, 61, 81, 84, 131-133, 154, 175, 196, 530, 853-855, 869, 887

Z

Zero

riferimenti generali, 26, 43, 74, 570, 735, 981, 1006, 1010

di una funzione, 579, 665

Indice dei nomi

Autori e studiosi citati nel testo italiano.

A

Abel, 82, 293, 304, 305, 307, 518, 969, 978

Adams, 17

Ado, 1012

Ahmes, 32

Al-Biruni, 75

Al-Khwarizmi, 49, 50, 75, 292

Aleksandrov, A. D., 14, 278, 436, 487, 811

Aleksandrov, P. S., 748, 863, 890, 913, 914, 916-918, 998

Alexander, 913, 914

Andronov, 360, 383, 385, 910

Apollonio, 46, 47, 56

Archimede, 27, 46, 47, 57, 77, 81, 144, 742

Aristotele, 77

B

Banach, 69

Beltrami, 826, 827, 835, 867, 885

Bernoulli, 60, 82, 336, 416, 634, 657, 660, 930

Bernštejn, 68, 611, 613, 614, 625, 634, 646, 652, 654, 665, 666

Bertrand, 577

Bessel, 190

Betti, 906, 907, 913

Bézout, 293

Birkhoff, 339, 383, 809, 913

Blaschke, 489

Bogoljubov, 383, 622

Boltjanskij, 913, 918

Bolyai, 66, 813, 817, 999

Bolzano, 68, 317, 318, 728, 746

Bonnet, 477

Borel, 68, 634

Boyle, 88, 270, 397, 849

Brahe, 212

Briggs, 50

Brouwer, 912, 913, 918

Bunjakovskij, 777, 926, 927

C

Cantor, 64, 68, 75, 730, 740, 742-744, 746, 748, 749, 751, 760, 912

Čaplygin, 535, 537, 538

Cardano, 293, 296, 298, 299

Cartan, 876, 988, 1013

Cartesio, 47, 50, 55, 57, 69, 75, 80, 197, 198, 201, 204, 207, 234, 292, 293, 323-326, 341, 901

Cauchy, 60, 68, 82, 98, 100, 147, 525, 528, 532, 538, 541, 546, 548, 549, 551, 552, 554, 567, 568, 728, 741, 742, 777, 926, 927, 950

Cavalieri, 57

Cayley, 846, 983

Čebotarëv, 308, 309, 1012, 1016, 1019

Čebyšëv, 68, 71, 82, 152, 439, 573, 576-578, 581, 584, 593, 611-614, 633, 634, 640, 646-652, 655, 661, 662, 665, 666, 931

Čech, 489

Clairaut, 234, 438

Cohn-Vossen, 490, 1020

Copernico, 17, 77, 196

Coulomb, 61

Courant, 73, 191, 433

Čudakov, 579

D

Dandelin, 341

Darboux, 439, 475

Darwin, 817

De la Vallée-Poussin, 634

De Moivre, 609

Dedekind, 68, 75, 574, 740-742, 1018

Delone, 196, 291

Desargues, 839

Diofanto, 47, 48

Dirichlet, 91, 432, 573, 579, 657-659, 753, 755, 760

Dürer, 839

D'Alembert, 180, 293, 309, 310, 316, 318-320, 416, 813, 930

E

Efimov, 476

Egorov, 753

Einstein, 277, 817, 867, 877, 881-883

Eratostene, 47, 575

Euclide, 18, 27, 32, 33, 37, 46, 47, 50, 63, 75, 572, 574, 576, 579, 598, 811-813, 815, 816, 824, 825, 833, 834, 839, 880

Eulero, 60, 63, 82, 108, 172, 228, 233, 234, 244, 292, 293, 365-367, 369, 416, 438, 458, 459, 461, 462, 467, 475, 482, 499, 502-504, 506-508, 511, 514, 521, 522, 573, 575, 576, 578, 580, 589, 634, 657, 670, 678, 696-698, 727, 901, 903, 907, 911, 930, 1004

F

Faber, 641, 666

Faddeev, 761

Fedorov, 34, 67, 71, 887, 952, 970, 975-977, 1020

Fejér, 660, 661, 663

Fermat, 196, 292, 341, 594, 1015

Ferrari, 50, 298

Finikov, 489

Finsler, 66, 876, 887

Fok, 883

Fourier, 325, 409, 416, 419, 634, 654, 655, 657-660, 662-664, 666, 686, 758, 864, 946

Frenet, 439

Fresnel, 693

Fubini, 489

G

Galerkin, 419-421, 432, 669, 691

Galois, 67, 293, 302, 305-309, 339, 340, 978-980, 1021

Gauss, 82, 293, 305, 308, 310, 439, 467, 469, 477, 478, 482, 492, 574, 576, 577, 646, 680, 813, 814, 817, 867, 877, 885

Gel'fand, 809, 920, 948, 988

Gel'fond, 574

Gibbs, 849

Gjunter, 431

Goldbach, 580, 587-589, 593

Graeffe, 341

Grassmann, 492, 846

Green, 935-937, 949

H

Hadamard, 578

Hamilton, 429, 431, 435, 492, 507, 508, 514, 999, 1001, 1003

Hardy, 73, 573, 596

Hausdorff, 758

Helmholtz, 347, 369, 858

Hertz, 17

Hilbert, 69, 82, 308, 309, 433, 490, 864, 884, 885, 889, 920, 924-926, 928-934, 940, 942, 943, 945, 947-950, 988, 1014, 1020

Hooke, 345, 936

Hopf, 433, 911, 913

Hurwitz, 329, 331, 341

J

Jackson, 634, 654, 664, 666

Jordan, 798, 808, 810

K

Kantorovič, 692

Keldyš, 515, 654, 748, 947

Keplero, 57, 77, 81, 196, 212, 343

Kirchhoff, 688

Kiselëv, 32

Klein, 82, 191, 309, 339, 809, 827, 831-833, 835, 841, 867, 882, 895, 911, 912, 987

Kolmogorov, 599, 625, 628, 634, 635, 660, 748, 759, 913, 948

Kolosov, 538

Korkin, 633

Krasnosel'skij, 916

Krylov, 339, 367, 383, 494, 667

Kuiper, 489, 493, 885, 889

Kummer, 574, 1015, 1016

Kurnakov, 849

Kuroš, 294, 998

L

Ladyženskaja, 386

Lagrange, 60, 63, 82, 133, 165-168, 194, 234, 235, 293, 302, 304, 306-308, 429, 574, 639, 646, 728, 813, 846, 881, 978, 984

Laguerre, 234

Lambert, 813, 815, 817

Laplace, 397-399, 401, 405, 408, 415-417, 423, 424, 428, 434, 538, 609, 623, 670, 686-688

Lappo-Danilevskij, 808

Laptev, 489

Lavrent'ev, 77, 366, 548, 634, 654, 748

Le Verrier, 17, 343

Lebedev, 692

Lebesgue, 68, 169, 194, 431, 634, 660, 664, 729, 749, 750, 753-755, 757, 758, 760, 928, 934, 948, 950

Lefschetz, 384, 913, 917

Legendre, 576, 577, 647, 813, 885, 931

Leibniz, 55, 57, 80, 81, 108, 144-149, 194, 292, 438, 483, 683, 691, 817

Leonardo da Vinci, 77, 839, 896

Leray, 913, 916

Lie, 294, 913, 987, 988, 998, 1008, 1012, 1013, 1021

Liouville, 305, 359

Littlewood, 573

Ljusternik, 486, 913, 914, 916

Lobačevskij, 18, 32, 66, 285, 293, 336, 338, 339, 341, 811, 814-821, 823-827, 829, 831-835, 838, 839, 842, 856, 862, 864, 867, 868, 873, 877, 880, 881, 883, 885-887, 889, 890, 892, 895-897, 912, 914, 918, 987, 1019

Lomonosov, 857

Lorentz, 253, 277, 278, 280-285, 882, 889

Lusin, 68, 635, 748, 749, 753, 754, 760

Lyapunov, 60, 68, 82, 377-380, 383, 385, 419, 614, 748

M

Mach, 880

Mal'cev, 951, 988

Mandel'štam, 383, 910

Mardžanišvili, 570

Mariotte, 88, 270, 397, 849

Markov, 82, 383, 613, 614, 625, 628, 631, 633, 652, 723

Markuševič, 918

Maxwell, 17, 858

Mendeleev, 652

Men'šov, 635

Mergeljan, 654

Meusnier, 460, 461, 463, 467, 482, 492, 493

Michelson, 277, 278

Minding, 439, 475, 485, 827

Minkowski, 487, 574, 881

Möbius, 274, 895-897, 902

Molin, 1010, 1011, 1017

Monge, 438, 485, 491, 492

Morozov, 988

Morse, 486, 916

Muschelišvili, 538

N

Najmark, 988

Nash, 489, 493

Nemyckij, 383, 910, 916

Nepero, 50, 55, 75

Neumann, John von, 725

Newton, 37, 43, 48-50, 55, 57, 59, 60, 63, 75, 80, 81, 97, 108, 140, 144-149, 152, 194, 199, 207, 208, 212, 221, 234, 292, 293, 325, 334-336, 343, 345, 392, 393, 397, 438, 522, 621, 633, 639, 817

Nikol'skij, 77, 632

Novikov, 748, 994

O

Odhner, 683, 691

Ostrogradskij, 60, 82, 153, 174-176, 194, 390, 419, 507-509, 514

P

Papaleksi, 383

Pascal, 82, 196, 683

Peterson, 439, 477, 482

Petrovskij, 288, 342, 625

Picard, 557, 568

Pitagora, 32, 36, 40, 55, 199, 210, 255, 443, 444, 868, 885, 922, 923, 928, 933, 934

Planck, 625

Plateau, 465

Pogorelov, 488, 884

Poincaré, 60, 68, 82, 277, 377, 382, 383, 486, 577, 817, 832, 835, 880, 882, 886, 900, 907, 909, 912, 913

Poisson, 397-399, 408, 417, 419

Poncelet, 839, 886

Pontrjagin, 383, 913, 917, 918, 952, 988, 1020

Popov, 17

Postnikov, 570, 913

R

Ramanujan, 573

Raševskij, 876

Riccati, 359

Riemann, 66, 82, 103, 288, 525, 528, 532, 538, 541, 543, 545, 546, 548, 558, 564-566, 569, 573, 578, 729, 752, 754, 755, 760, 844, 867, 868, 873, 874, 876, 877, 881, 885, 912, 987

Riesz, 69, 948

Ritz, 432, 669

Rolle, 322, 324

S

Saccheri, 813, 815, 817, 885

Schauder, 916

Schläfli, 850

Schrödinger, 431

Schwarz, 777, 810

Selberg, 597

Serre, 913

Serret, 293

Simpson, 642-646, 666, 670, 676, 677

Sitnikov, 914, 918

Smirnov, 194, 809, 916

Šnirel'man, 486, 913

Sobolev, 386, 429, 432

Stečkin, 728

Stepanov, 383, 416, 622, 910

Stevin, 50

Sturm, 293, 325, 326, 332, 341

Šubnikov, 978

Suslin, 748

Sylvester, 325, 577

T

Tartaglia, 50, 293, 296

Taylor, 125, 136-140, 163, 164, 186-188, 193, 371, 449, 462, 463, 475, 481, 517, 555, 558, 560, 561, 568, 632, 633, 638, 639, 666, 673, 691, 948

Tichonov, 916

Tolomeo, 46, 48, 50, 75

Torricelli, 82

Trefftz, 432

U

Uryson, 912-914, 916, 918

V

Vagner, 876

Van der Waerden, 73, 294, 809

Vavilov, 192

Veblen, 913

Viète, 48, 50, 292, 294, 300, 301, 489, 978, 1021

Vinogradov, 71, 573, 574, 578-580, 587, 589, 593, 596, 598, 635, 1016

Voronoj, 574, 853, 887, 1016

W

Wallace, 817

Wallis, 813, 885

Weierstrass, 68, 75, 82, 110, 317, 318, 340, 633, 634, 652-654, 741, 742, 746

Weyl, 1013, 1020

Wilczynski, 489

Wilson, 465

Z

Zolotarëv, 293, 574, 633, 1016

Žukovskij, 17, 314, 535, 537, 538