Università degli Studi di Bologna
Facoltà di Scienze matematiche, fisiche e naturali

Corso di Laurea in Matematica

Lo sviluppo storico del concetto di funzione e le origini della teoria delle distribuzioni

Candidato:
Lorenzo Pantieri

Relatore:
Prof. Ermanno Lanconelli

Anno Accademico 1999–2000

Lorenzo Pantieri

Lo sviluppo storico del concetto di funzione
e le origini della teoria delle distribuzioni

Copyright © 2012

lorenzo.pantieri@gmail.com

La citazione è un utile sostituto dell'arguzia.

— Oscar Wilde

Dedicato alla cara memoria di Gabriella Pantieri.

1944-1998

INDICE

INTRODUZIONE IX

1 FUNZIONI 1
- 1.1 Origini del concetto di funzione 1
- 1.2 Nozione euleriana di funzione 2
- 1.3 Problema della corda vibrante 5
- 1.4 Mutamenti nel concetto di funzione 11
- 1.5 Lagrange e il tentativo di rigorizzazione 14
- 1.6 Fourier e la fisica-matematica 17
- 1.7 Fondamenti dell'analisi in Cauchy 20
- 1.8 Funzione di Dirichlet 24
- 1.9 Aritmetizzazione dell'analisi 28

2 DISTRIBUZIONI 37
- 2.1 Dalle funzioni alle distribuzioni 37
- 2.2 Heaviside e il calcolo operazionale 40
- 2.3 Delta di Dirac 43
- 2.4 Sobolev e la fondazione delle distribuzioni 46
- 2.5 Teoria delle distribuzioni di Schwartz 52

CONCLUSIONI 61

A ANALISI NON-STANDARD 63

BIBLIOGRAFIA 69

ELENCO DELLE FIGURE

Figura 1 Il problema della corda vibrante 6
Figura 2 Una «curva discontinua» nel senso euleriano 8
Figura 3 Un esempio di «funzione» dato da Fourier 18
Figura 4 La «funzione impulsiva unitaria» di Heaviside 41
Figura 5 Le funzioni impulsive secondo Van der Pol 42
Figura 6 La Delta di Dirac 56
Figura 7 La retta iperreale 67

L'architettura della matematica è come quella di una città,
i cui sobborghi non cessano di crescere, talvolta in modo un po' caotico,
mentre il centro viene periodicamente ricostruito,
ogni volta seguendo un piano più chiaro e un ordinamento più maestoso,
distruggendo i vecchi quartieri con il loro dedalo di viuzze,
per lanciare verso la periferia dei viali sempre più diretti, più larghi e più comodi.

— Nicholas Bourbaki

RINGRAZIAMENTI

Ringrazio il mio Relatore di tesi di laurea in Matematica, Prof. Ermanno Lanconelli, dell'Università di Bologna. La sua alta professionalità ed esperienza, unite alla sua gentilezza e disponibilità, sono state indispensabili per la riuscita del mio lavoro di ricerca.

E un grazie davvero speciale ai miei genitori, che mi hanno sempre sostenuto con affetto e generosità.

Bologna, 17 novembre 2000 L. P.

INTRODUZIONE

In questo lavoro di tesi mi sono proposto di valutare, attraverso la ricostruzione storica del percorso delle conoscenze matematiche che ne consentirono l'attuazione, le *motivazioni* che determinarono lo sviluppo storico del concetto di funzione e che diedero origine alla teoria delle distribuzioni.

Obiettivo specifico del lavoro è mostrare come le ricerche dei matematici impegnati nel problema della fondazione del concetto di funzione e della nozione di distribuzione non siano avvenute sulla spinta di una sempre maggiore generalizzazione dei risultati o a causa di «astratte» questioni di rigore: di fatto, alla generalità e alla formulazione di nuovi standard di rigore si perviene quando «i vecchi criteri non permettono una risposta adeguata alle domande che vengono dalla pratica matematica o addirittura da problemi in certo senso esterni alla matematica che, trattati matematicamente, impongono mutamenti del quadro teorico [Bottazzini, 1981, p. 13]». La matematica appare sostanzialmente come *un'attività di proposta e di soluzione di problemi*, teorici o pratici, puri o applicati.

Così non è un caso che la fisica-matematica e più in generale la matematica applicata siano state un motore formidabile per lo sviluppo della matematica pura. E non è neppure un caso che nuovi criteri di rigore si presentino il più delle volte nella formulazione delle *definizioni* anziché nelle *dimostrazioni*: il momento della definizione rientra infatti nell'assetto complessivo in cui una teoria si organizza ed è, da un punto di vista storico, *conseguente* alla scoperta matematica vera e propria.

Questo è un punto delicato, la cui comprensione è decisiva se si vuole intendere lo sviluppo *reale* della matematica e non vederlo deformato dalle lenti «razionali» della critica contemporanea. La storia della matematica evidenzia che lo sviluppo della disciplina avviene a piccoli passi, mediante risultati che spesso provengono da varie direzioni e mette in luce l'importanza della presenza di motivazioni adeguate per lo sviluppo delle idee matematiche, che non scaturiscono perfette dalla mente di qualche genio, ma sono il frutto di molti anni di studio, di tentativi e anche di errori. In questo modo, è possibile accorgersi della differenza che intercorre fra la matematica come sistema organicamente strutturato e come disciplina *in fieri*: spesso gli obiettivi concettuali raggiunti in un dato momento storico non sono che un punto di partenza, e risulta evidente che molte lacune devono essere colmate o che le estensioni veramente importanti devono ancora essere create.

L'esposizione del lavoro è articolata come segue:

NEL PRIMO CAPITOLO viene discussa l'evoluzione storica del concetto di funzione e viene mostrato come la moderna formulazione del concetto di funzione rappresenti il *punto di arrivo* di una discussione che ha accompagnato la storia del calcolo infinitesimale dalle sue origini nella seconda metà del Seicento fino al programma di «aritmetizzazione dell'analisi» nella seconda metà dell'Ottocento.

NEL SECONDO CAPITOLO vengono analizzate le origini della teoria delle distribuzioni, cardine della matematica del XX secolo, e viene evidenziato come l'introduzione delle distribuzioni sia dovuta non alla mera

tendenza ad estendere il più possibile le nozioni dell'analisi, ma a *problemi del tutto concreti di matematica e di fisica-matematica*, che rendono necessaria un'estensione del concetto di funzione.

NELLE CONCLUSIONI vengono tratte le conclusioni della ricerca e viene presentato un bilancio dello studio svolto e dei risultati ottenuti.

IN APPENDICE, infine, viene dato un cenno all'analisi non-standard, che ha riformulato l'analisi moderna attraverso l'arricchimento del campo dei numeri reali ed il recupero di buona parte dell'impostazione originale di Leibniz e del concetto di infinitesimo.

1 FUNZIONI

Nel fascicolo dei risultati del primo libro degli *Eléments de mathématique* di Bourbaki [1939], dedicato alle strutture fondamentali dell'analisi, si legge questa definizione di funzione:

> Siano E e F due insiemi distinti o no. Una relazione fra una variabile x di E e una variabile y di F è detta *relazione funzionale in* y, o *relazione funzionale di* E *verso* F, se qualunque sia $x \in E$, esiste un elemento y di F, e uno solo, che stia nella relazione considerata con x. Si dà il nome di *funzione* all'operazione che associa così ad ogni elemento $x \in E$, l'elemento y di F che si trova nella relazione data con x; si dice che y è il *valore* della funzione per l'elemento x e che la funzione è *determinata* dalla relazione funzionale considerata. Due relazioni funzionali equivalenti determinano la stessa funzione [Bourbaki, 1939, p. 6].

Il concetto di funzione appare qui definitivamente basato sulla teoria degli insiemi: una relazione funzionale tra due insiemi è definita come un particolare sottoinsieme del prodotto cartesiano $E \times F$. In maniera analoga si trova definita la nozione di funzione nelle opere più recenti di Dieudonné [1969] o di Kolmogorov e Fomin [1974].

La definizione in termini insiemistici è l'esito di una discussione che ha accompagnato la storia dell'analisi dalle origini del calcolo infinitesimale nella seconda metà del Seicento. Nella formulazione oggi usuale del concetto di funzione come applicazione fra insiemi astratti va però perduta una delle idee centrali, di natura fisica, che originariamente stavano alla base dell'analisi, l'idea di studiare matematicamente il movimento dei corpi e dunque la «variazione delle grandezze».

È questo il passo che separa nettamente la matematica classica, tipicamente la geometria della Grecia antica, tramandata dagli arabi e «riscoperta» nei testi di Euclide e Archimede nel sedicesimo secolo, dalla matematica moderna. Nel 1870 il matematico tedesco Hankel (1839-1873) scriveva:

> La matematica moderna data dal momento in cui Descartes va oltre la trattazione puramente algebrica delle equazioni per indagare le variazioni delle grandezze che un'espressione algebrica subisce quando una grandezza generica in essa contenuta percorre una successione continua di valori [Hankel, 1870, p. 63].

Si tratta di un passo decisivo, che trova nel calcolo infinitesimale il significato più compiuto e dà conto della radicale differenza che intercorre fra la moderna analisi e l'«algebra geometrica» degli antichi, ed anche fra la cinematica di Galilei e la dinamica di Newton.

1.1 ORIGINI DEL CONCETTO DI FUNZIONE

Il termine «funzione» si trova per la prima volta in Leibniz e Johann Bernoulli. In un manoscritto del 1673, dal titolo *Methodus tangentium inversa, seu de functionibus*, Leibniz usò il termine «funzione» per denotare una qualsiasi quantità variabile da punto a punto di una curva come, ad esempio, la

lunghezza della tangente o della normale. Della curva veniva detto che era data da un'equazione. Nella sua *Historia* (1714), Leibniz adopera la parola «funzione» per denotare, in generale, quantità che dipendono da una variabile. Leibniz introdusse anche i termini «costante», «variabile» e «parametro», usando quest'ultimo con riferimento a una famiglia di curve. Lavorando con le funzioni, Johann Bernoulli parla, a partire dal 1697, di quantità formate, in una maniera qualsiasi, da variabili e da costanti; con le parole «una maniera qualsiasi» egli intendeva riferirsi sia alle espressioni algebriche sia a quelle trascendenti. Bernoulli adottò nel 1698 la frase leibniziana «funzione di x» per indicare questa quantità.[1]

Quanto alle notazioni, Johann Bernoulli scriveva X o ξ per denotare una funzione generale di x. Nel 1718, però, adottò la notazione φx. Leibniz approvò questo cambiamento, ma propose anche di adoperare x^1 e x^2 per indicare delle funzioni di x quando erano in gioco più funzioni.

Se formalmente il termine «funzione» si trova per la prima volta in Leibniz e Johann Bernoulli, è con Newton che emerge, sulla base di motivazioni fisiche, la stretta relazione fra il concetto di funzione e quello di *variazione*. Per Newton il movimento dei corpi è al centro della ricerca e l'impostazione analitica è rivolta alla considerazione delle quantità variabili in funzione di un parametro temporale, denominate fin dalle prime origini del calcolo infinitesimale (cioè a partire dal 1655) «fluenti», e delle loro derivate prime, denominate «flussioni». Il metodo delle flussioni offre gli strumenti matematici per descrivere le variazioni delle «grandezze fluenti» (cioè le funzioni), cose queste che «hanno veramente luogo in natura»:

> Io considero qui le quantità matematiche non come costituite da parti molto piccole, ma come descritte da un moto continuo. Le linee sono descritte, e quindi generate, non dalla giustapposizione delle loro parti, ma dal moto continuo dei punti [...]. Questa genesi ha effettivamente luogo in natura e può essere vista quotidianamente nel moto dei corpi [Newton, *Tractatus de quadratura curvarum*, 1676].

Il concetto di funzione acquistò immediatamente un ruolo centrale nelle ricerche sul calcolo infinitesimale e, come vedremo, fu specialmente in relazione ai tentativi volti a chiarificare e ad estendere questa nozione che si affermò l'esigenza della rigorizzazione dell'analisi ed ebbe origine la teoria delle distribuzioni.

1.2 NOZIONE EULERIANA DI FUNZIONE

I trattati euleriani *Introductio in analysin infinitorum* [1748], *Institutiones calculi differentialis* [1755] e *Institutiones calculi integralis* (1768-1770) rappresentano il punto di arrivo della speculazione analitica del periodo, lungo circa un secolo, che va dal 1655, anno al quale risalgono le prime ricerche newtoniane sul metodo delle flussioni, fino alla metà del Settecento. Al contempo, essi rappresentano il punto di partenza dell'analisi matematica moderna che, attraverso i contributi di autori quali Cauchy e Weierstrass, giungerà alla sistemazione concettuale dell'inizio del XX secolo. Lo sviluppo del calcolo infinitesimale ha seguito diverse vie in Inghilterra e nel continente europeo:

1 Secondo alcuni storici, quello di dipendenza funzionale, come altri concetti della matematica, appare già nei testi più antichi, e c'è chi ha parlato di «istinto per la funzionalità» presso i babilonesi [Bottazzini, 1981, p. 50].

> Rigidamente vincolata alla tradizione newtoniana (e al suo infelice formalismo), la matematica inglese nel Settecento sarà incapace di cogliere la straordinaria quantità di risultati e di tecniche che la maggiore flessibilità e fecondità della tradizione leibniziana aveva assicurato ai matematici continentali [Bottazzini, 1981, p. 21].

Di fatto, l'atteggiamento e il formalismo leibniziano appaiono dominanti nell'analisi matematica del Settecento e, coerentemente con essi, il concetto di dipendenza funzionale fra quantità variabili è quello che Euler (1707-1783), sicuramente il matematico più originale e fecondo del secolo, esprime con le seguenti parole:

> Una funzione di quantità variabili è un'espressione analitica composta in modo qualunque da quelle quantità e da numeri o quantità costanti [Euler, 1748, vol. 8, p. 4].

È questa la definizione che si legge in apertura del primo volume dell'*Introductio in analysin infinitorum*, un trattato standard della matematica settecentesca. Qui è già del tutto assente ogni riferimento fisico al movimento dei corpi e il concetto di funzione viene espresso in termini puramente formali come combinazione di quantità (variabili e costanti) e di segni d'operazione.[2]

Col termine «espressione analitica» si intende, per Euler, un'espressione composta da grandezze simboliche e numeri mediante *operazioni algebriche* (addizione, sottrazione, moltiplicazione e divisione, elevamento a potenza ed estrazione di radice, «alle quali bisogna aggiungere ancora la risoluzione delle equazioni») *oppure trascendenti*, quali l'esponenziale e il logaritmo «e innumerevoli altre che ci fornisce il calcolo integrale».

A questa distinzione è correlata, per Euler, quella tra funzioni algebriche e trascendenti: le prime sono ottenibili mediante un numero finito di operazioni elementari (le equazioni algebriche sono in linea di principio risolubili algebricamente, è opinione di Euler) mentre per le seconde egli ritiene senz'altro che si possano sviluppare in serie (o comunque mediante un numero infinito di operazioni elementari), senza porsi il problema né della dimostrazione né della legittimità di tali estensioni. Così, nel quarto capitolo dell'*Introductio*, considera come la maniera più generale di esprimere una funzione una serie del tipo

$$A + Bz + Cz^2 + Dz^3 + \dots \tag{1}$$

«Se qualcuno ne dubita — scrive Euler — il dubbio sarà tolto dallo sviluppo di ciascuna funzione [Euler, 1748, p. 16]».

In ogni caso, non essendo in grado di dimostrare la sviluppabilità di una funzione qualunque $f(z)$ in serie di potenze, Euler lascia aperta la possibilità di considerare esponenti qualunque per la z nello sviluppo (1), possibilità che esprime nei termini seguenti:

> Affinché questa spiegazione valga nella maniera più estesa, oltre alle potenze di z che hanno esponenti interi positivi si debbono ammettere anche potenze a esponente qualunque. Così non vi sarà più alcun dubbio che ogni funzione della stessa z si possa in questo modo trasformare in una espressione infinita, denotando α, β, γ, ecc. numeri qualunque [Euler, 1748, p. 74].

2 Ad Euler si deve, fra l'altro, l'introduzione (nel 1734) della notazione $f(x)$ per indicare una funzione di x.

Dunque una funzione qualunque di z è rappresentabile dalla somma (infinita o no)

$$Az^{\alpha} + Bz^{\beta} + Cz^{\gamma} + \dots \tag{2}$$

ed è proprio questa la proprietà che Euler ha in mente e utilizza negli sviluppi di funzioni in serie, frequenti nell'*Introductio*.

Del resto, come ha osservato Bottazzini [1981, p. 23], le funzioni usate in analisi al tempo di Euler erano, nella grande maggioranza, analitiche nel senso oggi usuale, salvo al più punti «eccezionali» del dominio di definizione e, in singoli casi, si potevano presentare esponenti frazionari o negativi nelle potenze dello sviluppo. Euler osserva sì che la rappresentazione (2) può venir meno in questi punti «eccezionali», ma sostanzialmente si limita alla considerazione di funzioni algebriche ed estende in generale le loro proprietà alle funzioni trascendenti.

> Poiché si era in grado di ottenere direttamente per le funzioni algebriche lo sviluppo in serie secondo potenze di un incremento, il quoziente differenziale e l'integrale, si riteneva non solo giusto ammettere l'esistenza di tali serie, del quoziente differenziale e dell'integrale, in maniera del tutto generale per tutte le funzioni, ma anche non veniva assolutamente in mente che allo scopo occorresse un'esplicita affermazione, sia questa oggi un assioma o un teorema; così appariva del tutto ragionevole ed evidente questa estensione delle proprietà delle funzioni algebriche alle trascendenti, estensione apparentemente legittimata dalla visualizzazione geometrica delle curve rappresentanti le funzioni; ed esempi in cui funzioni propriamente analitiche mostravano singolarità che erano essenzialmente diverse da quelle delle funzioni algebriche passavano del tutto inosservati [Hankel, 1870, p. 64].

Così Hankel [1870], che sostiene il proprio argomento con esempi di funzioni come $\sin \frac{1}{x}$, $e^{1/x}$, $\frac{1}{1+e^{1/x}}$, ecc. considerate nel punto $x = 0$.

Certamente non è difficile, dal nostro punto di vista, rilevare l'inadeguatezza della classificazione euleriana di funzioni algebriche e trascendenti. Infatti il carattere algebrico o trascendente di una funzione non si lascia rivelare dal particolare tipo di «espressione analitica» impiegata nella definizione: così serie di potenze possono definire funzioni sia algebriche, come è il caso di

$$y(x) = 1 + \frac{x}{2} - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \dots = \sqrt{1+x} \quad \text{per } |x| < 1 \tag{3}$$

sia trascendenti, come

$$y(x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \dots = \log(1+x) \quad \text{per } |x| < 1 \tag{4}$$

Il modo di ragionare sostanzialmente formale di Euler, in cui ha un peso decisivo l'analogia supposta esistente tra il finito e l'infinito, è tipico dell'epoca. Tuttavia alcune procedure, oggi ritenute illegittime, non sono non rigorose per Euler: lo sono rispetto ai nostri criteri di rigore, passati attraverso il filtro di duecento anni di sviluppo dell'analisi. Questo è un punto delicato, la cui comprensione è decisiva se si vuole intendere lo sviluppo *reale* della matematica e non vederlo deformato dalle lenti «razionali» della critica contemporanea.

> Il rigore in matematica è anch'esso un concetto storico e dunque in divenire [...]. Appellarsi all'esigenza di rigore nello spiegare lo sviluppo della matematica sembra in realtà un discorso circolare: di fatto, alla formulazione di nuovi standard di rigore si perviene quando i vecchi criteri non permettono una risposta adeguata alle domande che vengono dalla pratica matematica o addirittura da problemi in certo senso esterni alla matematica che, trattati matematicamente, impongono mutamenti del quadro teorico [Bottazzini, 1981, p. 13].

Così non è un caso che la fisica matematica e più in generale la matematica applicata siano state un motore formidabile per lo sviluppo della matematica pura. E non è neppure un caso che nuovi criteri di rigore si presentino il più delle volte nella formulazione delle *definizioni* anziché nelle *dimostrazioni*: il momento della definizione rientra infatti nell'assetto complessivo in cui una teoria si organizza ed è, da un punto di vista storico, *conseguente* alla scoperta matematica vera e propria.

1.3 PROBLEMA DELLA CORDA VIBRANTE

La discussione intorno al concetto di funzione diventa centrale verso la metà del Settecento in una questione di carattere *fisico-matematico*, quella di studiare le vibrazioni di una corda omogenea non soggetta ad alcuna forza esterna, posta in un piano e fissata alle estremità. Intorno alla soluzione di questo problema, che ebbe un impatto decisivo anche sulla matematica «pura», si alimentò una lunga e vivace polemica fra i più grandi matematici del secolo, inizialmente D'Alembert e Euler, e poi Daniel Bernoulli e J. L. Lagrange. Si tratta di una vicenda di cui si trovano echi ancora all'inizio dell'Ottocento in J. B. Fourier. La discussione si accese intorno a un lavoro di D'Alembert (1717-1783) del 1747, ma pubblicato due anni dopo, che rappresenta il primo tentativo coronato da successo di integrare le equazioni differenziali alle derivate parziali che si ottengono descrivendo matematicamente le infinite forme assunte da una corda tesa posta in vibrazione in un piano.

Come nel caso delle equazioni differenziali ordinarie, il cui studio nacque dalle ricerche in numerose classi di problemi fisici (meccanica, gravitazione, teoria dell'elasticità, ecc.), i matematici non crearono di proposito la teoria delle equazioni differenziali alle derivate parziali. Essi continuarono a indagare gli stessi problemi fisici che li avevano condotti alle equazioni differenziali ordinarie e, a mano a mano che coglievano con maggior sicurezza i principi fisici che stanno alla base dei fenomeni, formularono enunciati matematici che fanno ora parte della teoria delle equazioni alle derivate parziali. Così, mentre lo spostamento di una corda vibrante era stato studiato separatamente come funzione del tempo e come funzione della distanza di un punto da una delle sue estremità, lo studio dello spostamento come funzione di entrambe le variabili e il tentativo di capire tutti i possibili moti condusse a un'equazione alle derivate parziali.

Nei primi approcci la corda vibrante era considerata un «rosario»: si supponeva cioè che essa contenesse n pesi uguali, posti a distanza uguale, uniti l'uno all'altro da pezzi di filo pensato privo di peso, flessibile e inestensibile. Per trattare la corda continua si supponeva che il numero dei pesi diventasse infinito mentre la loro massa diminuiva in modo che la massa totale del numero crescente dei singoli «grani» tendesse alla massa della corda conti-

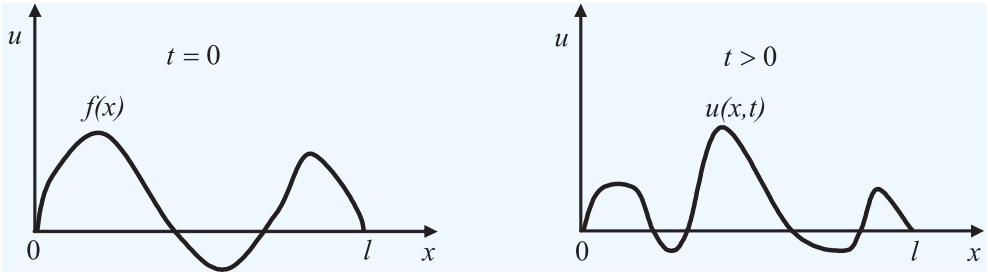


Figura 1.: La discussione intorno al concetto di funzione diventa centrale, verso la metà del Settecento, nella risoluzione del «problema della corda vibrante».

nua. C'erano delle difficoltà matematiche nel passaggio al limite, ma queste erano considerate sottigliezze e venivano ignorate.

Il caso di un numero discreto di masse era stato trattato da Johann Bernoulli nel 1727, che aveva ottenuto per lo spostamento u_k della k-esima massa, l'equazione alle differenze finite

$$\frac{d^2u_k}{dt^2} = \nu^2(u_{k+1} - 2u_k + u_{k-1}) \quad \text{per } k = 1, 2, \ldots, n-1 \tag{5}$$

dove ν^2 dipende dalla tensione della corda (pensata costante durante le vibrazioni), dalla massa totale degli n corpi e dalla loro distanza reciproca. Bernoulli risolve l'equazione e passa poi a trattare il caso di una corda continua: la sua conclusione è che la corda ad ogni istante t assume una forma sinusoidale, la cui equazione egli ottiene integrando l'equazione differenziale $d^2u/dx^2 = -ku$: un risultato, questo, che era stato ottenuto qualche tempo prima da Taylor (1685-1731).

Quando D'Alembert comincia a interessarsi alla questione (1747), il suo obiettivo è di far vedere che una corda in vibrazione assume infinite altre forme oltre a quella sinusoidale. Introducendo un sistema di riferimento cartesiano e considerando al posto degli u_k una funzione $u = u(x, t)$ definita per $0 \leqslant x \leqslant l$ e $t \geqslant 0$, D'Alembert ottiene, al posto dell'equazione alle differenze finite (5), la seguente

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \nu^2 \frac{\partial^2 u(x,t)}{\partial x^2} \tag{6}$$

valida per $0 < x < l$ e $t > 0$. Compare così per la prima volta quella che viene oggi detta «equazione unidimensionale delle onde».

Poiché la corda è fissata alle estremità $x = 0$ e $x = l$, la soluzione deve soddisfare le condizioni al contorno

$$u(0,t) = 0 \qquad u(l,t) = 0 \tag{7}$$

per $t \geqslant 0$. Supponiamo che, per $t = 0$, la corda abbia la configurazione indicata dalla funzione $f(x)$ e che la velocità dei suoi punti sia data dalla funzione $g(x)$. Queste condizioni sono espresse matematicamente da

$$u(x,0) = f(x) \quad \left.\frac{\partial u(x,t)}{\partial t}\right|_{t=0} = g(x) \tag{8}$$

per $0 \leqslant x \leqslant l$. Poiché i capi della corda in $x = 0$ e $x = l$ sono mantenuti fermi, a quota o, si deve assumere, conseguentemente,

$$f(0) = f(l) = g(0) = g(l) = 0 \tag{9}$$

Con un procedimento analogo a quello che userà nel suo saggio sull'idrodinamica intitolato *Essai d'une nouvelle théorie de la resistance des fluides* (1752), D'Alembert riuscì ad integrare l'equazione (6), ottenendo la formula

$$u(x,t) = \frac{1}{2}[f(x+vt) + f(x-vt)] + \frac{1}{2v}\int_{x-vt}^{x+vt} g(t)\,dt \tag{10}$$

D'Alembert sottolinea con vigore il fatto che la $f(x)$ è soggetta alla «legge di continuità», espressione con cui D'Alembert intende che $f(x)$ è data da un'unica espressione analitica, ed è altresì due volte differenziabile, dal momento che è $u(x,0) = f(x)$ e che la $u(x,t)$ deve soddisfare l'equazione (6).

L'anno successivo Euler intervenne sulla questione con la memoria *Sur la vibration des cordes*. Tecnicamente la soluzione di Euler non si discosta da quella di D'Alembert, alla quale tuttavia egli dichiara di voler aggiungere «qualche osservazione abbastanza interessante nell'applicazione delle formule generali». Egli considera infatti la stessa equazione (6) studiata da D'Alembert e sottolinea l'intento di ricercare la massima generalità possibile della soluzione, «affinché la figura iniziale della corda possa essere tracciata arbitrariamente [Euler, 1750b, p. 76]». Essa può essere data infatti da una curva «sia regolare contenuta in una certa equazione, sia irregolare o meccanica». È proprio sul carattere della $f(x)$, cioè della funzione che descrive la posizione iniziale della corda, che le posizioni di D'Alembert e Euler differiscono sostanzialmente.

Le divergenze vengono presto alla luce in una risposta di D'Alembert [1752] all'articolo di Euler: l'oggetto del contendere è essenzialmente il concetto di funzione, cui ci si riconduce immediatamente quando si tenta di precisare la natura degli oggetti matematici «soluzione» dell'equazione differenziale (6).

> Non si può, mi sembra, esprimere analiticamente y in maniera più generale che supponendola una funzione di t e di x. Con questa supposizione, però, si trova la soluzione del problema solo nel caso in cui le diverse figure della corda vibrante possono essere comprese in una stessa equazione [D'Alembert, 1752, p. 358].

Conclusione questa che è coerente con le idee allora dominanti in analisi sulla continuità, idee che lo stesso Euler aveva proprio in quel periodo autorevolmente suffragato nell'*Introductio*. Nel secondo volume di questo trattato, Euler, dopo aver introdotto un sistema cartesiano di riferimento nel piano, scrive:

> Benché si possano descrivere meccanicamente diverse linee curve mediante il movimento continuo di un punto, che ci fa vedere la curva nel suo complesso, noi qui le consideriamo principalmente come il risultato di funzioni, essendo questa maniera di considerarle più analitica, più generale e più adatta al calcolo. Così una funzione qualunque di x darà una certa linea retta o curva, da cui segue che, reciprocamente, si potranno mettere in relazione le linee curve con le funzioni. Di conseguenza, la natura di una linea curva sarà determinata da una funzione di x [...]. Da questa concezione delle linee curve discende naturalmente la loro divisione in *continue* e in *discontinue* o *miste*. La linea curva continua è quella la cui natura è espressa *da una sola funzione determinata di* x [corsivo nostro]. Se però la linea curva è composta da differenti parti determinate da più funzioni di x, di modo che una parte sia il risultato di una funzione e un'altra sia il risultato di una seconda funzione, noi chiamiamo queste specie di linee curve *discontinue*, o *miste* e

> *irregolari*, giacché esse non sono formate secondo una legge costante e sono composte di porzioni di differenti curve continue.
>
> In geometria si ha a che fare principalmente con curve continue e nel seguito si mostrerà che le curve, che sono descritte meccanicamente da un movimento uniforme secondo una certa legge costante, possono essere espresse da un'unica funzione e di conseguenza sono delle curve continue [Euler, 1748, vol. 9, p. 10-11].

Euler illustra il proprio argomento la figura 2. La classificazione delle curve operata da Euler rimase standard per un lungo periodo e si ritrova ancora all'inizio dell'Ottocento. Dietro la vaga terminologia dell'epoca (oltre a quelle viste, si parlava di curve «totalmente discontinue», «tracciate con un libero movimento della mano», «che obbediscono alla legge di continuità», «meccaniche», «arbitrarie», «algebriche», «trascendenti», ecc.) si riconoscono, dal nostro punto di vista, due tipi di funzioni: quelle *analitiche* (che sono le «curve continue» secondo Euler) e quelle *continue regolari a tratti* (le «discontinue» o «miste» nella terminologia euleriana), mentre non vengono prese in esame le funzioni che *oggi* sono dette *discontinue*.

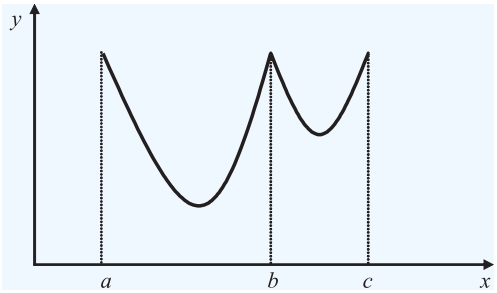


Figura 2.: Un esempio di «curva discontinua» nella terminologia euleriana: la curva è «composta da differenti parti determinate da più funzioni di x [Euler, 1748, vol. 9, p. 4]».

In questo quadro teorico si capisce come D'Alembert, nel difendere la propria soluzione del problema della corda vibrante, affermi che «in ogni altro caso il problema non potrà almeno con il mio metodo, e mi domando se pure non sia superiore alle forze dell'analisi conosciuta [D'Alembert, 1752, p. 358]».

È però lo stesso Euler a spingersi oltre, motivato dalla natura fisica del problema. È del tutto ragionevole supporre, infatti, che la corda, allorché viene posta in vibrazione, assuma una forma iniziale arbitraria (tipicamente, se si «pizzica» la corda per farla vibrare, la configurazione iniziale può essere descritta da una funzione che presenta un punto angoloso e che quindi non è ivi derivabile):

> La prima vibrazione dipende soltanto da noi (*de notre bon plaisir* giacché si può, prima di lasciar la corda, darle una figura qualunque; il che fa sì che il movimento vibratorio della stessa corda possa variare all'infinito, a seconda che le si dia la tale o la talaltra forma all'inizio del movimento [Euler, 1750a, p. 217].

Coerentemente, Euler conclude che

> [...] le diverse parti di questa curva non sono dunque legate fra loro da alcuna legge di continuità [di analiticità] e sono tenute insieme solo dalla descrizione [del fenomeno]; [...] la sola considerazione del tratto di curva [compresa fra $x = 0$ e $x = l$] è sufficiente a farci conoscere il movimento della corda, senza assoggettarlo al calcolo [Euler, 1750a, p. 217].

La fisica del problema della corda vibrante è la molla che spinge Euler a enunciare il suo nuovo concetto di funzione. La polemica finì così per fornire occasione di discussione sul concetto di funzione e sulle funzioni ammissibili in analisi, da cui dipendeva la generalità della soluzione trovata. Come osserva Bottazzini [1981, p. 30], D'Alembert e Euler usano sì lo stesso termine «funzione», ma intendono cose diverse: il primo giunge alla conclusione che «per non andar contro le regole dell'analisi» la soluzione è sensata quando y sia un'espressione analitica di x e t, Euler al contrario non richiede alcuna «legge di continuità» per la curva e dunque alcuna espressione analitica per la y.

Consapevole delle implicazioni connesse al suo atteggiamento, Euler scriverà a D'Alembert che «la considerazione di tali funzioni non soggette ad alcuna legge di continuità [di analiticità] apre davanti a noi un campo dell'analisi interamente nuovo».

Nelle *Institutiones calculi differentalis* [1750] Euler darà una definizione assai generale di funzione (anche se *in pratica* si limiterà alle funzioni analitiche):

> Se delle quantità dipendono da altre in modo tale che dalle mutazioni di queste anche le prime subiscano delle variazioni, esse si usano chiamare funzioni di queste. Questa denominazione ha un'estensione molto ampia e comprende in sé tutti i modi con i quali una quantità si può determinare per mezzo di altre. Se dunque x rappresenta una quantità variabile, allora tutte le quantità che dipendono da x in un modo qualunque o possono determinarsi per mezzo di essa, sono chiamate funzioni di essa [Euler, 1750a, p. 4].

Non sembra immotivato concludere che la grande generalità di questa definizione sia stata suggerita a Euler proprio dalla vicenda della corda vibrante: se gli oggetti matematici che risolvono l'equazione (6) sono funzioni di x e t, allora non è certamente il caso, per Euler, di pensare per essi a una qualche sorta di «esprimibilità analitica», e se nel suo trattato egli si limita a queste ultime è perché a queste si riducono le funzioni più usuali; il che tuttavia non implica il limitarsi *a priori* ad esse come ai soli oggetti denominabili «funzioni».

Recentemente Demidov [1977] ha fatto notare che la divergenza di opinioni fra D'Alembert e Euler corrisponde anche a due modi diversi di integrare l'equazione delle corde: mentre D'Alembert ottiene, con la (10), la soluzione oggi detta «classica», Euler sembra muoversi nella direzione di considerare soluzioni «deboli» o «generalizzate» dell'equazione (6).

Nel 1753 si inserisce nella discussione il figlio di Johann Bernoulli, Daniel Bernoulli (1700-1782). Il suo approccio è caratterizzato dalle sue precedenti ricerche fisiche (di acustica) e proprio sulla base di motivazioni di natura fisica (la sovrapposizione e la composizione delle onde) egli asserisce che *tutte* le possibili curve iniziali possono essere rappresentate nella forma

$$f(x) = \sum_{n=1}^{\infty} a_n \sin \frac{n\pi x}{l} \tag{11}$$

Di conseguenza, dice, *tutti* i moti successivi saranno del tipo

$$u(x,t) = \sum_{n=1}^{\infty} a_n \sin \frac{n\pi x}{l} \cos \frac{n\pi c t}{l} \tag{12}$$

(nel caso in cui la velocità iniziale sia nulla). In questa forma, è convinzione di Bernoulli, si possono porre le soluzioni di D'Alembert e di Euler.

Tutti i possibili moti della corda vibrante sono dunque periodici nel tempo. Ogni moto corrispondente a una qualsiasi curva iniziale è una somma di «modi» periodici sinusuoidali e la combinazione ha la frequenza del «modo fondamentale»; i periodi dei singoli modi sono la metà, un terzo, e così via del periodo fondamentale. Tuttavia, Daniel Bernoulli non fornisce nessun ragionamento matematico per sostenere le sue affermazioni, ma si basa soltanto sulla fisica. Nel lavoro del 1753 egli afferma:

> La mia conclusione è che tutti i corpi sonori includono un'infinità di suoni con una corrispondente infinità di vibrazioni regolari [...]. Ciascun tipo [cioè ciascun modo fondamentale relativo ad una curva iniziale] si moltiplica un numero infinito di volte per accordare a ciascun intervallo un numero infinito di curve tali che ciascun punto inizi e concluda nello stesso istante queste vibrazioni [...]. Osserviamo poi che la corda non può compiere vibrazioni soltanto secondo la prima figura [la fondamentale] o la seconda [il «secondo armonico»] o la terza e così via all'infinito, ma può effettuare una combinazione di queste vibrazioni in tutte le combinazioni possibili [Bernoulli, 1755, p. 198].

Euler [1750] sottolinea sì l'importanza delle osservazioni di Bernoulli sulla natura fisica del problema, ma dichiara inaccettabile la conclusione che la (12) possa rappresentare una arbitraria funzione «discontinua», come richiede l'integrazione della (6); inoltre, osserva Euler, una funzione somma di una serie trigonometrica è periodica, e se la $f(x)$ non ha questa proprietà (il che accade in generale) non si può rappresentarla con una tale serie. Anche se Euler sottolinea con particolare vigore l'argomento della periodicità, l'obiezione non è calzante, se si tiene presente che «l'analisi si riferisce solo all'intervallo $0 \leqslant x \leqslant l$ dell'asse x su cui è tesa la corda: che cosa succede fuori da questo intervallo è irrilevante per la corda vibrante e quindi per la matematica impiegata a descriverla [Bottazzini, 1981, p. 32]».

Si tratta di un'incomprensione non banale, che si rivela piuttosto come un aspetto delle contraddizioni tra vecchia e nuova teoria delle funzioni, entrambe presenti in Euler, protagonista di questa trasformazione. Infatti, nella concezione classica, che si trova ancora in Euler, una funzione è pensata associata alla totalità del dominio naturale su cui è definita, mentre la distinzione (euleriana, si badi) in funzioni «continue» e «discontinue» già prelude all'idea di considerare funzioni definite «a tratti», mediante espressioni analitiche diverse in diverse parti del dominio, che queste espressioni si «raccordino» con continuità o no. Questo è un passo decisivo, se si considera che è qui anticipata *in nuce* l'idea oggi usuale di *dominio* della funzione. Euler non coglie dunque tutte le implicazioni presenti nella sua concezione delle funzioni discontinue e rimane ancorato alla sua originaria impostazione.

Anche D'Alembert (1761) si oppose alla soluzione proposta da Bernoulli, facendo proprio l'argomento di Euler sull'impossibilità di rappresentare mediante una serie trigonometrica una generica funzione. Ha scritto a questo proposito Lebesgue [1906]:

> Se l'affermazione di Bernoulli fosse stata esatta, occorreva che la somma di una serie trigonometrica potesse uguagliare una funzione lineare [come si ottiene nel caso di una figura poligonale assunta come posizione iniziale della corda] in un intervallo e un'altra funzione lineare in un altro intervallo; o, se si vuole, bisognava che due espressioni analitiche fossero uguali in un intervallo e diverse in un altro [Lebesgue, 1906, p. 21].

E in nota lo stesso Lebesgue aggiungeva:

> Siccome si ammetteva che due espressioni analitiche uguali in un intervallo sono uguali dappertutto, si ammetteva che era sufficiente dare una funzione con una definizione analitica in un intervallo, per quanto piccolo, perché essa fosse per ciò stesso determinata anche in tutto il suo dominio d'esistenza. Da qui il nome di *funzioni continue* dato da Euler a queste funzioni. È solo dopo Cauchy che il termine *funzione continua* ha acquisito il senso attuale. La proprietà che Euler pensava di riconoscere nelle sue funzioni continue è quella in realtà che caratterizza le funzioni analitiche di una variabile complessa [Lebesgue, 1906, p. 21].

Nel 1759 fece il suo ingresso nella controversia anche Lagrange, che in un primo tempo accettò il punto di vista di D'Alembert (ammettendo una funzione almeno due volte differenziabile per la posizione iniziale della corda), e successivamente difese la posizione di Euler (secondo cui la curva iniziale può essere «discontinua»). Il dibattito continuò a infuriare per tutti gli anni '60 e '70. Euler, D'Alembert e Lagrange continuarono a negare che la somma di una serie trigonometrica possa rappresentare una arbitraria funzione «discontinua» (nel senso euleriano). Questo punto fondamentale non fu risolto fino a quando non venne affrontato da Fourier.

Se si giudica sulla base delle conoscenze del tempo, tutti, in un certo senso, avevano ragione. D'Alembert, seguendo una tradizione stabilita fin dai tempi di Leibniz, insisteva sul fatto che le funzioni devono essere analitiche, in modo che ogni problema non risolvibile mediante esse sia insolubile *tout court*. Euler, D'Alembert e Lagrange non si rendevano conto che, data una funzione arbitraria nell'intervallo $0 \leqslant x \leqslant l$, è possibile ripeterla in ogni intervallo $[nl, (n+1)l]$, con n intero, in modo che la funzione diventi periodica. Naturalmente, una tale funzione periodica può non essere rappresentabile da un'unica espressione analitica. Euler, D'Alembert e Lagrange avevano ragione nel ritenere che non tutte le funzioni possono essere rappresentate da serie di Fourier, e al contempo nel credere che la curva iniziale può essere molto generale. Essa, infatti, non è necessariamente né analitica né periodica. Bernoulli adottò la posizione corretta su basi fisiche, ma non fu in grado di suffragarla con il ragionamento matematico.

Un altro punto che non venne chiarito fu la giustificazione del perché un'equazione alle derivate parziali con coefficienti analitici (nel caso in questione, costanti) possa avere una soluzione non analitica. Nel caso delle equazioni differenziali ordinarie, se i coefficienti sono analitici anche le soluzioni devono esserlo. Questo non è invece vero per equazioni alle derivate parziali. Anche se Euler aveva ragione quando diceva (e insisteva su questo fatto) che le soluzioni «angolose» sono ammissibili, la determinazione delle singolarità ammissibili per le soluzioni delle equazioni alle derivate parziali si ebbe solo nel corso del XIX secolo.

1.4 MUTAMENTI NEL CONCETTO DI FUNZIONE

Uno degli esiti più interessanti della polemica sulla corda vibrante fu dunque quello di concentrare l'attenzione dei matematici sulla definizione euleriana di funzione, allora standard, e in particolare di funzione continua e discontinua, e di cercare di comprendere, rispetto a quella, quale fosse la natura degli oggetti matematici ottenuti dall'integrazione di equazioni differenziali alle derivate parziali. Così nel 1787 l'Accademia di Pietroburgo

bandiva un premio per la miglior soluzione di determinare «se le funzioni arbitrarie cui si perviene mediante l'integrazione di equazioni a tre o più variabili rappresentino delle curve o superfici qualunque, sia algebriche o trascendenti, sia meccaniche, discontinue, o generate da un movimento arbitrario della mano; o se queste funzioni comprendano soltanto delle curve continue rappresentate da un'equazione algebrica o trascendente».

Il premio fu vinto da una memoria di L. F. Arbogast (1759-1803), di cui l'aspetto più interessante è la distinzione che egli fa tra funzioni «discontinue» e funzioni «discontigue». Secondo le parole di Arbogast [1791],

> [...] la legge di continuità consiste nel fatto che una quantità non può passare da uno stato ad un altro senza passare attraverso tutti gli stati intermedi che sono soggetti alla stessa legge.
>
> Le funzioni algebriche sono considerate continue, poiché i differenti valori di queste funzioni dipendono nella stessa maniera da quelli della variabile, e supponendo che la variabile cresca continuamente, la funzione subirà variazioni in modo corrispondente, ma non passerà da un valore ad un altro senza passare attraverso tutti i valori intermedi. Quindi l'ordinata y di una curva algebrica, quando l'ascissa x varia, non può passare bruscamente da un valore ad un altro; non ci può essere un salto fra un'ordinata e un'altra che differisce da essa di una quantità prefissata, ma tutti i successivi valori di y devono essere collegati fra loro da una stessa legge [...]. Questa continuità può essere vanificata in due modi:
>
> (*i*) La funzione può cambiare la sua forma, vale a dire la legge secondo cui la funzione dipende dalla variabile può cambiare del tutto. Una curva formata dall'unione di alcune porzioni di curve differenti è di questo tipo. Non è neppure necessario che la funzione y debba essere espressa da un'equazione in un certo intervallo della variabile; essa può continuamente cambiare la sua forma e la linea che la rappresenta, al posto di essere l'unione di curve regolari, può essere tale che in ognuno dei suoi punti diventi una curva differente, in altre parole può essere interamente irregolare e non seguire alcuna legge per ogni intervallo comunque piccolo. Tale sarebbe una curva tracciata a caso dal libero movimento della mano. Questo tipo di curve non può essere rappresentato né da una né da più equazioni algebriche o trascendenti.
>
> (*ii*) La legge di continuità viene meno anche quando le differenti parti di una curva non si congiungono fra loro [Arbogast, 1791, p. 9].

Queste ultime Arbogast le chiama «curve discontigue», cioè, in termini moderni, funzioni *discontinue*. Le funzioni che entrano in gioco nell'integrazione delle equazioni differenziali alle derivate parziali sono, per Arbogast, di entrambi i tipi; per esempio, dice Arbogast, nell'equazione della corda vibrante si può pensare ad un «salto» nei valori di $\partial^2 y/\partial x^2$, purché lo stesso salto avvenga per i valori di $\partial^2 y/\partial t^2$.

La distinzione di Arbogast tra funzioni «discontinue» e «discontigue» tende a distinguere tra funzioni continue, anche se definite in diversi intervalli da diverse dipendenze funzionali, e funzioni discontinue nel senso moderno. Inoltre, nel caso (*i*), trattando delle discontinuità, Arbogast sembra lasciare intravedere una definizione puramente formale di funzione, anche se si affretta a precisare che ha in mente «curve tracciate a caso dal libero movimento della mano».

Le precisazioni di Arbogast alla definizione di Euler e l'introduzione di un nuovo concetto di discontinuità più vicino al senso moderno, si accompagnarono a un diffuso atteggiamento critico verso il concetto classico di funzione come espressione analitica composta di variabili e costanti.

Il primo che sembra aver colto la generalità della revisione euleriana del concetto di funzione è M. A. Condorcet (1734-1794), forse più noto come uomo politico e come riformatore nel campo dell'istruzione che come matematico.

Condorcet fu autore di un *Traité du calcul integral* pubblicato a Parigi nel 1765 e di una seconda opera (con lo stesso titolo) rimasta incompiuta. È proprio in questo secondo lavoro, di cui numerosi fogli a stampa dovettero circolare a Parigi ed essere noti fra i matematici, che Condorcet dà la seguente definizione di funzione:

> Suppongo di avere un certo numero di quantità x, y, x, ..., F e che ogni valore determinato di x, y, z, ..., F abbia uno o più valori determinati che corrispondono ad essi. Io dico allora che F è una funzione di x, y, z, ... Infine, so che, allorché x, y, z, ... saranno determinati, lo sarà anche F; anche se non conoscerò né la maniera di esprimere F mediante x, y, z, ... né la forma dell'equazione tra F e x, y, z, ... io saprò che F è funzione di x, y, z, ... [Condorcet, 1765].

È questo quanto Condorcet intende per funzione «analitica»: una funzione di natura del tutto arbitraria, dove il termine «analitico» sta a designare che lo studio di simili oggetti avviene nel contesto dell'analisi. Condorcet distingue inoltre tre diversi tipi di funzioni:

(*i*) funzioni di cui si conosce la forma (in altre parole, funzioni esplicite);

(*ii*) funzioni introdotte da equazioni tra F e x, y, z, ... (vale a dire funzioni implicite);

(*iii*) funzioni date mediante certe condizioni (per esempio equazioni differenziali).

Tuttavia, lo studio concreto delle funzioni è da Condorcet ricondotto (e in ciò egli fu influenzato da Lagrange, come vedremo) alla serie di Taylor, che viene assunta come fondamento del calcolo differenziale, e la determinazione dei coefficienti della serie porta Condorcet al concetto di «funzione differenziale», l'analogo delle «funzioni derivate» di Lagrange.

Fra coloro che ebbero per le mani i fogli inediti di questo trattato di Condorcet vi fu S. F. Lacroix (1765-1843), autore di numerosi trattati, fra cui un'opera sul calcolo differenziale e integrale che per tutta la prima metà dell'Ottocento conobbe una fama indiscussa, numerose riedizioni e traduzioni (anche in italiano, nel 1829). Nella prefazione della seconda edizione Lacroix [1810] dà un breve riassunto del lavoro di Condorcet e si appresta poi a precisare che cosa si intende per funzione. Dopo aver richiamato alcune definizioni classiche («funzione di una quantità è ogni potenza di questa quantità, oppure ogni espressione algebrica contenente quella quantità»), Lacroix scrive:

> Infine, delle nuove idee, portate dallo sviluppo dell'analisi, hanno dato luogo alla definizione seguente di funzione: ogni quantità il cui valore dipenda da una o più altre quantità è detta *funzione* di queste ultime sia che si sappia sia che si ignori attraverso quali operazioni occorra passare per risalire da queste alla prima [Lacroix, 1810, p. 2].

La definizione si richiama apertamente a quelle di Euler [1750] e di Condorcet. Tuttavia occorre sottolineare l'aspetto *nominale* di questa definizione, nel

senso che poi in pratica le funzioni studiate erano sostanzialmente quelle algebriche e trascendenti elementari.

Il tentativo più coerente di stabilire una teoria delle funzioni (e, nel tempo stesso, di fondare in maniera inequivoca il calcolo infinitesimale) fu operato da Lagrange con un trattato che in un certo senso rappresenta in modo emblematico lo stato dell'analisi e dei suoi fondamenti alla fine del Settecento.

1.5 LAGRANGE E IL TENTATIVO DI RIGORIZZAZIONE

Parallelamente all'espansione dei concetti e delle tecniche del calcolo infinitesimale vi furono dei tentativi di fornirgli i fondamenti mancanti. I matematici, infatti, oltre ad appoggiarsi notevolmente ai significati fisici e intuitivi, avevano generalmente in mente il modello fornito dalle funzioni polinomiali e razionali, ed estendevano a tutte le funzioni le loro proprietà: la continuità, l'esistenza di derivate e di integrali e la sviluppabilità in serie di potenze. Quando furono però obbligati, soprattutto in seguito alle ricerche sul problema della corda vibrante, ad estendere il concetto di funzione e a considerare anche funzioni non analitiche, non poterono più usare come guida le funzioni più semplici e dovettero procedere, in realtà, senza alcuna base sicura.

Lagrange, nella *Théorie des fonctions analytiques* del 1797, compì il primo, ambizioso tentativo di costruire le fondamenta del calcolo infinitesimale. Oggetto della *Théorie* sono la teoria delle funzioni e i princìpi del calcolo differenziale «liberati da ogni considerazione di infinitesimi, di quantità evanescenti, di limiti e flussioni, e ricondotti all'analisi algebrica di quantità finite», come scrive Lagrange nel sottotitolo dell'opera. Il trattato si apre con la seguente definizione di funzione:

> Si chiama funzione di una o più quantità ogni espressione del calcolo nella quale queste quantità entrano in maniera qualunque, insieme o no con altre quantità che si considerano come aventi dei valori dati e costanti, mentre le quantità della funzione possono assumere ogni valore possibile [Lagrange, 1797, p. 1].

Tale definizione, che appare ben più restrittiva di quella data circa quarant'anni prima da Euler nelle *Institutiones*, si richiama direttamente a Leibniz e Johann Bernoulli: «Essi l'hanno per primi usata in quest'accezione generale — osserva Lagrange — ed è oggi generalmente adottata». Si tratta di una valutazione discutibile, ma che per Lagrange assume il carattere di un indispensabile punto di partenza di tutta la teoria, poiché il passo immediatamente successivo è di far vedere che una funzione *qualunque* è sviluppabile in serie di potenze. Scrive infatti Lagrange:

> Consideriamo una funzione $f(x)$ di una variabile qualunque x. Se al posto di x si mette $x + h$, dove h è una quantità indeterminata qualunque, essa diventerà $f(x + h)$, e, mediante la teoria delle serie, si potrà sviluppare in una serie della forma $f(x) + ph + qh^2 + rh^3 + \ldots$, nella quale le quantità p, q, r, $\ldots$, coefficienti delle potenze di h, saranno delle nuove funzioni di x, derivate dalla funzione primitiva $f(x)$ e indipendenti dalla quantità h [...]. La formazione e il calcolo di queste diverse funzioni sono, a dire il vero, l'autentico oggetto dei nuovi tipi di calcolo, cioè del cosiddetto calcolo differenziale [Lagrange, 1797, p. 2].

Prima, ogni funzione è sviluppabile in serie di potenze di procedere a dimostrare che ogni funzione $f(x)$ può essere espressa nella forma

$$f(x+h) = f(x) + ph + qh^2 + rh^3 + sh^4 + \dots \tag{13}$$

dove i coefficienti p, q, r, s, ... dipendono in generale da x ma non da h, Lagrange vuole essere sicuro che un tale sviluppo in serie di potenze sia sempre possibile. Naturalmente, dice, questo fatto ci è noto da un gran numero di esempi familiari, ma ammette che ci sono dei casi eccezionali. Come casi eccezionali Lagrange ha in mente quelli in cui la funzione, o qualche sua derivata, diventa «infinita». Queste eccezioni capitano però soltanto in punti «isolati» e perciò, per Lagrange, non contano. Con analoga disinvoltura tratta poi una seconda difficoltà. Lagrange ed Euler accettavano senza discussione il fatto che uno sviluppo in serie contenente potenze intere e frazionarie di h fosse sempre possibile, ma Lagrange voleva eliminare la necessità di far ricorso a esponenti frazionari. Egli riteneva che questi potessero comparire soltanto quando $f(x)$ contiene dei radicali e anche questi vengono trascurati come casi eccezionali. A questo punto è perciò pronto a procedere con lo studio della (13).

Con un ragionamento alquanto involuto e puramente formale Lagrange conclude che $2q$ si ottiene da p allo stesso modo in cui da $f(x)$ si ottiene p e una conclusione simile vale per gli altri coefficienti r, s, ... della (13). Se si denota perciò p con $f'(x)$ e si indica con $f''(x)$ la funzione «derivata» da $f'(x)$ allo stesso modo in cui $f'(x)$ si deriva da $f(x)$, allora

$$p = f'(x) \quad q = \frac{1}{2!}f''(x) \quad r = \frac{1}{3!}f'''(x) \quad \dots \tag{14}$$

e perciò la (13) diventa:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f'(x) + \frac{h^3}{3!}f'''(x) + \dots \tag{15}$$

Lagrange conclude che «l'espressione finale ha il vantaggio di mostrare come i termini della serie dipendano l'uno dall'altro e, in particolare, come si possano formare tutte le funzioni derivate che compaiono nella serie, una volta calcolata la prima funzione derivata». Poco dopo aggiunge: «Per chi conosca i rudimenti del calcolo differenziale è chiaro che queste funzioni derivate coincidono con dy/dx, d^2y/dx^2, ...». Parole che ricordano in maniera impressionante quelle impiegate da Leibniz per presentare il calcolo differenziale nel 1684. «Chiunque conosca un po' di geometria saprà operare con il nuovo calcolo», aveva detto allora Leibniz. Il che, allora, era pura propaganda.

Lagrange deve ancora far vedere come si faccia a ricavare p, cioè $f'(x)$, da $f(x)$. Qui usa la (13) e, trascurando tutti i termini successivi al secondo, ottiene $f(x+h) - f(x) = ph$. Divide poi per h e conclude che $p = f'(x)$. In realtà, l'assunzione lagrangiana che ogni funzione possa essere sviluppata in serie di potenze è il punto debole di tutto lo schema. I criteri oggi noti per l'esistenza di un tale sviluppo coinvolgono l'esistenza delle derivate, che era proprio quello che Lagrange cercava di evitare.

Infine non si trova in Lagrange, conformemente allo spirito dell'epoca, alcun tentativo di assicurarsi della convergenza della serie (13). Del resto, proprio a proposito dell'uso delle serie in analisi, egli stesso aveva scritto diversi anni prima, al tempo della sua soluzione del problema della corda

vibrante, discutendo la soluzione in serie trigonometrica proposta da Daniel Bernoulli:

> Io mi domando se, ogni volta che in una formula algebrica si troverà per esempio una serie geometrica infinita come $1 + x + x^2 + x^3 + \ldots$, non si sarà in diritto di sostituirvi $\frac{1}{1-x}$, benché questa quantità non sia realmente uguale alla somma della serie proposta che supponendo l'ultimo termine x^∞ nullo. Mi sembra che non si potrebbe contestare l'esattezza di una simile sostituzione senza stravolgere i princìpi più comuni dell'analisi [Lagrange, *Œvres*, vol. 1, p. 323].

Come sarà chiaro ai matematici all'inizio dell'Ottocento, la sostituzione è però lecita solo quando $|x| < 1$ (che è quanto Lagrange intende dicendo, con il linguaggio del suo tempo, che x^∞ è nullo).

Fra i risultati più significativi della *Théorie des fonctions analytiques* c'è sicuramente l'espressione lagrangiana del resto nello sviluppo in serie di Taylor, che però non svolge alcun ruolo nelle dimostrazioni precedenti. Nonostante i difetti, l'approccio lagrangiano al calcolo infinitesimale incontrò per un certo periodo un grande favore. In seguito fu però abbandonato, proprio per le critiche sempre più precise e convincenti sull'uso delle serie senza opportuni criteri che ne assicurassero la convergenza, compito questo cui si accinsero i matematici dell'Ottocento ad iniziare da Abel e Cauchy. I criteri di rigore in analisi, sostenuti soprattutto da Cauchy, avevano reso l'«analisi algebrica» di Lagrange assai poco rigorosa agli occhi dei matematici, e insufficiente per la fondazione del calcolo infinitesimale.

Nonostante l'inadeguatezza del suo approccio, Lagrange contribuì, come già aveva fatto Euler, a separare i fondamenti del calcolo infinitesimale dalla geometria e dalla meccanica e nel fare ciò la sua influenza fu decisiva. Lagrange sottolinea in varie occasioni l'esigenza di un progressivo distacco dal riferimento geometrico nel trattare i problemi d'analisi e la necessità di autonomi criteri di coerenza per l'analisi. Queste esigenze caratterizzeranno fortemente lo sviluppo del calcolo infinitesimale dell'Ottocento.

Nel *Traité de Mécanique analytique* [1788], un testo fondamentale per più di una generazione di matematici, Lagrange sottolinea, non senza soddisfazione, che

> [...] in quest'opera non si troveranno affatto figure. I metodi che vi espongo non richiedono né costruzioni né ragionamenti geometrici o meccanici, ma soltanto delle *operazioni algebriche* [...] [corsivo nostro]. Quelli che amano l'analisi, vedranno con piacere la meccanica divenirne una nuova branca e mi saranno grati di averne così esteso il dominio [Lagrange, 1788, p. 2].

Analogo è il punto di vista nel trattare i fondamenti del calcolo e della teoria delle funzioni: il discorso si mantiene sul piano puramente algebrico, senza alcun ricorso a considerazioni intuitive.

In questo dunque Lagrange si può considerare un coerente e consapevole pioniere di una tendenza che diverrà prevalente da Cauchy in poi. Se il fondamento da lui individuato (l'«analisi algebrica delle quantità finite») si mostrerà inadeguato, nondimeno ci si muoverà su una strada simile, lontana dal riferimento all'evidenza geometrica e volta a trovare nell'ambito dell'analisi, ed esclusivamente in questo, il fondamento profondo, la «metafisica» del calcolo infinitesimale.

Ci si sta incamminando qui, almeno a livello di esigenza teorica, sulla strada che porterà, dopo settant'anni, alla cosiddetta «aritmetizzazione dell'analisi», ovvero alla costituzione dell'analisi sul fondamento fornito dall'aritmetica dei numeri naturali; una strada che vede, come tappa fondamentale, la specializzazione e la separazione all'interno della matematica di diverse branche di ricerca. Motivato da fattori di diversa natura, questo sarà il processo dominante nella matematica dell'Ottocento e ancora caratteristico della matematica contemporanea.

1.6 FOURIER E LA FISICA-MATEMATICA

L'opera di J. B. Fourier (1768-1830) impose un'ulteriore revisione al concetto di funzione e comportò un sostanziale ampliamento della classe di funzioni ammissibili in analisi. Nella *Théorie analytique de la chaleur* [1822], Fourier affrontò un problema fisico di grande interesse teorico e pratico: lo studio della natura e della propagazione del calore. Una questione che, fin dalla fine del Settecento, aveva attirato l'interesse di numerosi fisici e matematici.

Nell'integrare le equazioni differenziali ottenute, Fourier fece largo uso di serie trigonometriche, determinandone opportunamente i coefficienti (le «serie di Fourier»); come abbiamo visto, serie di questo tipo erano già note in matematica, ma «senza il *background* del fenomeno fisico tali questioni sarebbero rimaste delle curiosità matematiche; nell'opera di Fourier sviluppi in fisica-matematica e sviluppi in analisi erano intimamente connessi, fornendosi reciproci stimoli e reciproche motivazioni [Grattan-Guinness, 1972]».

Fourier considera vari casi di fenomeni di propagazione, da quella in una lamina di lunghezza indefinita, a quella in un anello circolare, in una sfera, in un prisma indefinito; in ciascun caso introduce opportune coordinate, dipendenti dalla simmetria del corpo, per semplificare l'equazione differenziale che rappresentava matematicamente il fenomeno e che in generale si scrive come

$$\frac{\partial v}{\partial t} = K\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right) \tag{16}$$

dove v è una funzione delle coordinate x, y, z e del tempo t, e K una costante fisica dipendente dalla conducibilità termica, dal calore specifico e dalla densità della sostanza di cui è composto il corpo.

Ciò che sollevò difficoltà fu non tanto l'equazione differenziale ottenuta da Fourier, quanto il metodo da lui adottato per integrarla. Il «nodo della questione» era che, sotto particolari condizioni, le soluzioni si potevano sviluppare in serie trigonometriche. Su questo punto, la memoria di Fourier incontrò obiezioni da parte dei commissari dell'Istituto di Francia (Lagrange, Laplace, Monge e Lacroix), cui era stata sottoposta per un giudizio.[3]

«Siccome questi risultati sembrano discostarsi dalle ordinarie conseguenze del calcolo — osservava egli a questo proposito — è necessario esaminarli con cura e interpretarli nel loro vero significato». Per esempio, l'equazione

$$y(u) = \cos(u) - \frac{1}{3}\cos(3u) + \frac{1}{5}\cos(5u) - \frac{1}{7}\cos(7u) + \dots \tag{17}$$

3 Monge, con ogni probabilità, giudicava favorevolmente il lavoro di Fourier; lo stesso si poteva dire per Lacroix e Laplace, mentre diverso, e determinante, fu il parere di Lagrange [Bottazzini, 1981, p. 81].

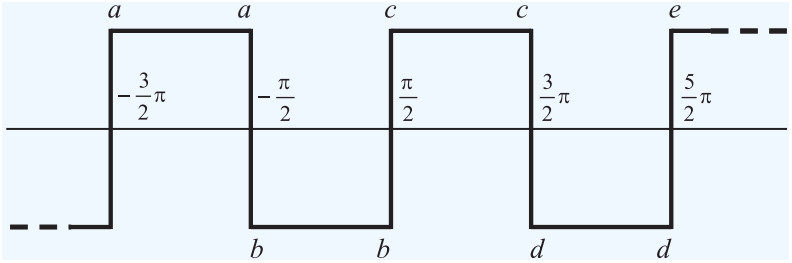


Figura 3.: Un esempio di «funzione» somma di una serie trigonometrica dato da Fourier nella *Théorie analytique de la chaleur*. Per Fourier le perpendicolari $ab, cb, cd, ed, \ldots$ fanno anch'esse parte del grafico della funzione [Fourier, 1822].

rappresenta nel piano u, y

> [...] una linea composta da parti separate aa, bb, cc, dd, ...di cui ciascuna è parallela all'asse [delle ascisse] e uguale a una semicirconferenza. Queste parallele sono situate alternativamente al di sopra e al di sotto dell'asse a distanza 1 e *sono unite dalle perpendicolari* ab, cb, cd, ed, *...che fanno anch'esse parte della linea* [corsivo nostro] [Fourier, 1822, p. 158].

Per farsi un'idea esatta della funzione definita dalla (17), Fourier suggeriva di considerare in un primo tempo solo un numero finito di termini della serie. In quest'ultimo caso, si ottiene

> [...] una curva che passa alternativamente ad di sopra e al di sotto dell'asse u, tagliandolo tutte le volte che l'ascissa u diventa uguale a $\pm\frac{1}{2}\pi$, $\pm\frac{3}{2}\pi$, $\pm\frac{5}{2}\pi$, ... Quanto più il numero dei termini dell'equazione aumenta, tanto più la curva in questione tende a confondersi con la linea precedente, composta di rette parallele e rette perpendicolari [Fourier, 1822, p. 158].

Il grafico che ha in mente Fourier (ma non lo disegna) è riportato nella figura 3. Dal punto di vista moderno, certamente un tale grafico non rappresenta una funzione, a meno che non si pensi a una funzione discontinua, immaginando il grafico senza i segmenti di perpendicolare ab, cb, cd, ed, ..., che invece Fourier espressamente considera come parte integrante della curva. In queste considerazioni è implicita la questione importante e delicata della continuità di una funzione, un concetto attorno al quale si intricano le concezioni dei matematici all'inizio dell'Ottocento. Fourier ha a questo riguardo un atteggiamento incerto, oscillante fra l'accettazione del punto di vista euleriano e l'intuizione di una concezione più moderna; anche la descrizione per così dire «genetica» della «spezzata» sopra disegnata testimonia questa sua ambiguità concettuale. Scrive Fourier:

> Risulta dalle mie ricerche che le funzioni arbitrarie anche discontinue possono sempre essere rappresentate da sviluppi in seno o coseno di archi multipli, conclusione che il celebre Euler ha sempre respinto [...]. Gli sviluppi in discorso hanno questo in comune con le equazioni differenziali alle derivate parziali, che essi possono esprimere la proprietà delle funzioni interamente arbitrarie e discontinue; è per questo che si presentano in maniera naturale per l'integrazione di queste ultime equazioni [Fourier, 1822, p. 185].

Che cosa intende qui Fourier per «discontinuità» di una funzione? Sorprendentemente, il primo esempio di «linea discontinua» è proprio la linea

aa, bb, cc, ... della figura 3 a fronte, presentata senza far parola dei segmenti di perpendicolare ab, cb, cd, ... Ciò corrisponde alla «moderna» definizione di continuità (e di discontinuità) che darà Cauchy nel 1821. Immediatamente dopo, però, egli (coerentemente con l'antica concezione euleriana) dichiara «discontinue» linee «composte da archi di parabola e segmenti di retta».

Fourier non enunciò mai le condizioni che deve soddisfare una funzione per poter essere sviluppata in serie trigonometrica. Tuttavia, la sua convinzione che ciò fosse *sempre* possibile è espressa durante tutto il corso della *Théorie*. Egli dice anche [Fourier, 1822, p. 196] che le sue serie sono sempre convergenti qualunque sia $f(x)$, che le si assegni o no un'espressione analitica. La convinzione di Fourier che ogni funzione possa essere sviluppata in serie trigonometrica si basava su considerazioni geometriche. Egli non ne diede mai una dimostrazione completa: nel libro porta alcune deboli argomentazioni e nella discussione finale di questo punto dà un abbozzo di dimostrazione.

Le ricerche di Fourier, oltre a far progredire la teoria delle equazioni alle derivate parziali, imposero una revisione del concetto di funzione. Fourier è esplicito nell'affermare che la sua serie può rappresentare funzioni non necessariamente definite da un'unica espressione analitica e *non necessariamente continue*. Nella *Théorie analytique de la chaleur* dice:

> In generale la funzione $f(x)$ rappresenta una successione di valori o di ordinate ciascuna delle quali è arbitraria. [...] Noi non supponiamo che queste ordinate siano soggette a una legge comune; esse si succedono l'una all'altra in maniera qualsiasi [Fourier, 1822, p. 197].

Anche se, in effetti, Fourier trattava soltanto funzioni dotate di un numero finito di discontinuità in ogni intervallo limitato, la sua opera segnò la definitiva rottura con la concezione delle funzioni analitiche come prototipo di tutte le funzioni. Poiché le proprietà delle funzioni analitiche non potevano più essere estese a tutte le funzioni, sorse il problema di individuare, al di là del riferimento grafico intuitivo, il significato preciso da dare ai concetti di funzione, di continuità, di derivabilità, di integrabilità e così via. Sono problemi, questi, che si presentano quando si affrontano, come fa Fourier, problemi concreti e non solo «astratte» questioni di rigore.

È importante osservare come Fourier sottolinei ripetutamente che le sue ricerche non avvengono sulla spinta di una sempre maggiore generalizzazione dei risultati: la generalità è un argomento indotto dai problemi affrontati. Questo approccio alla matematica resterà largamente dominante nella matematica francese del secolo scorso: l'impostazione che Fourier dà alla sua *Théorie* si può interpretare come l'enunciazione e la prima realizzazione di «un programma di ricerca» al compimento del quale in Francia lavoreranno (più o meno consapevolmente) i matematici di più generazioni.

Elementi distintivi di questo programma si possono individuare nella forte motivazione fisica alla ricerca matematica, nel rifiuto del formalismo euleriano e lagrangiano e, più generalmente, nella concezione della scienza come elemento di progresso civile. Se il primo argomento non è certamente nuovo, e gran parte della matematica settecentesca si motiva inizialmente con questioni di natura fisica, dall'origine del calcolo infinitesimale di Newton al problema della corda vibrante, alla teoria del potenziale gravitazionale e alla propagazione delle onde, nuovo e cosciente è il ruolo svolto dalla fisica nello sviluppo delle matematiche.

Le stesse affermazioni sulla rappresentabilità in serie, che sembrano svolgersi su un terreno esclusivamente analitico, sono da Fourier motivate proprio dalla necessità di risolvere problemi fisici; senza queste, egli dice, non si potrebbe fare un passo in avanti nelle sue ricerche. Quest'ultimo argomento chiarisce anche il carattere della matematica usata da Fourier,

> [...] del tutto nuovo rispetto, per esempio, a quello «algebrico» proprio della *Théorie analytique* di Lagrange. Mi pare, anzi, che questo sia il contrasto di fondo tra i due, che emerge invece su un problema specifico (quello della rappresentabilità delle funzioni in serie trigonometriche).
>
> Il concetto di funzione, in Fourier, non è rivestibile con l'abito stretto imposto da Lagrange, né tantomeno la sua idea di «analisi» è quella lagrangiana di studio *algebrico* di quantità finite. Il «rigore» lagrangiano (di cui non a caso la trattazione di Fourier è accusata di essere carente) viene sostituito da un'altra concezione, non *formale*. Ad esempio, per quanto riguarda le serie (un punto cruciale nella matematica dell'inizio del secolo scorso) è rigorosa per Fourier non già la trattazione (se pure formalmente ineccepibile) delle funzioni mediante serie di Taylor, ma un *calcolo effettivo* dei primi n termini, che mostra già come debba essere il limite per n infinito, e poi una verifica dei risultati sui dati (matematici o fisici) da cui si è partiti. Egli si serve sì della serie di Taylor nei primi passi della ricerca sulla rappresentabilità in serie trigonometriche di funzioni arbitrarie, ma se ne libera ben presto e anzi non è soddisfatto che quando riesce a esibire *direttamente* lo sviluppo cercato [Bottazzini, 1981, p. 80].

La matematica deve trovare nella realtà esterna (fisica) stimoli e motivazioni: questa «concretezza» (e anche modernità, dopo più di centocinquant'anni) è l'indicazione che emerge con chiarezza dall'opera di Fourier.

1.7 FONDAMENTI DELL'ANALISI IN CAUCHY

«Voglio dedicarmi ora con tutte le mie forze a introdurre un po' di chiarezza nella prodigiosa oscurità che innegabilmente c'è oggigiorno in analisi», aveva scritto Abel (1802-1829) in una delle sue lettere sotto l'evidente influenza dell'opera di Cauchy (1789-1857). Questi, una delle figure dominanti della matematica moderna, si era impegnato in un sistematico lavoro di ricostruzione rigorosa dei concetti fondamentali del calcolo differenziale.

Il progetto culturale di Cauchy appare chiaro: egli voleva svincolare l'analisi dai procedimenti poco rigorosi (e talvolta nettamente scorretti) derivanti dall'applicazione teoricamente non fondata di metodi di comodo, che spesso avevano l'unico scopo di giustificare risultati magari solo verificati in qualche caso particolare o praticamente utili [Bottazzini, 1981, p. 100].

Fin dalle pagine introduttive del *Cours d'analyse algébrique* [1821] Cauchy esprimeva con grande vigore le concezioni che lo avevano guidato nello stabilire i fondamenti dell'analisi matematica: «Quanto ai metodi — scriveva — ho cercato di dar loro tutto il rigore che si esige in geometria in modo da non ricorrere mai ad argomenti tratti dalla generalità dell'algebra [Cauchy, 1821, p. 2]». Questi ultimi erano stati gli argomenti preferiti da Lagrange. «È vero che, per mantenermi costantemente fedele a questi princìpi — egli aggiungeva — mi sono visto costretto ad ammettere diverse proposizioni che sembreranno forse un po' dure a prima vista [Cauchy, 1821, p. 4]». La prima delle proposizioni «un po' dure» da ammettere era che «una serie

divergente non ha somma», cosa che oggi appare del tutto ragionevole, ma che all'epoca andava contro una tradizione consolidata, della quale abbiamo visto Lagrange farsi autorevole portavoce.

Le serie giocano in analisi un ruolo decisivo, ed estremo deve essere il rigore nel trattarle, anche a costo di drastiche riduzioni dell'estensione delle formule usate.

> Così, prima di effettuare la somma di una qualsiasi serie, ho dovuto esaminare in quali casi le serie possono essere sommate, o, in altri termini, quali sono le condizioni per la loro convergenza; e, a questo riguardo, ho stabilito delle regole generali che mi sembrano meritare qualche attenzione [Cauchy, 1821, p. 5].

La *Théorie des fonctions analytiques* era stata una delle opere più studiate dal giovane Cauchy. Se egli ora concordava con Lagrange sulla necessità di fondare in modo rigoroso il calcolo infinitesimale, senza limitarsi a giustificarne i metodi con il successo nelle applicazioni o il ricorso a considerazioni intuitive, ne prendeva tuttavia apertamente le distanze quando si trattava di individuarne i fondamenti: gli argomenti di natura algebrica erano liquidati come «delle induzioni adatte a far talvolta presentire la verità, ma che poco s'accordano con l'esattezza tanto vantata delle scienze matematiche [Cauchy, 1821, p. 3]».

Lo strumento che nelle mani di Cauchy diventò la chiave di volta di tutte le costruzioni dell'analisi fu il concetto di limite, che egli così definiva:

> Allorché i valori successivamente assunti da una stessa variabile si avvicinano indefinitamente a un valore fissato, sì da differirne alla fine tanto poco quanto si vorrà, quest'ultima quantità è chiamata *il limite* di tutte le altre [Cauchy, 1821, p. 3].

Mediante esso Cauchy definiva sia la tanto controversa nozione di infinitesimo, una variabile «che ha zero come limite», sia quella di infinito, una variabile i cui successivi valori numerici «crescono sempre più, in modo da superare ogni numero dato».

Si tratta di definizioni assolutamente rigorose? No, se esaminate con la sensibilità matematica dei nostri giorni. Il confronto con la letteratura matematica precedente evidenzia però una radicale inversione di tendenza: l'impostazione di Cauchy riflette ampiamente l'impostazione della moderna analisi, nella quale il limite deve essere considerato il concetto fondamentale. Nel primo capitolo si trova immediatamente la nozione di funzione di una variabile reale:

> Allorché delle quantità variabili sono legate fra loro in modo tale che, dato il valore di una, si possa ricavare il valore di tutte le altre, [queste], espresse per mezzo della variabile indipendente, sono chiamate funzioni di questa variabile [Cauchy, 1821, p. 19].

Dopo aver definito in maniera analoga le funzioni di più variabili e distinto tra funzioni esplicite e funzioni implicite, Cauchy dà la seguente definizione di continuità di una funzione:

> La funzione $f(x)$ resterà continua rispetto a x fra due limiti dati, se, entro questi limiti, un incremento infinitesimo della variabile produce sempre un incremento infinitesimo della funzione stessa [Cauchy, 1821, p. 34].

La definizione di funzione che dà Cauchy appare del tutto svincolata dalla possibilità di fornire un'«espressione analitica» della variabile dipendente (una posizione, questa, contraria al punto di vista di Lagrange, per esempio) così come la continuità, intesa da Cauchy, comprende anche funzioni non dovunque derivabili nel loro dominio di definizione; come già in Bolzano (1781-1848), anche in Cauchy si trova espressa nettamente l'idea che la funzione vada considerata «entro due limiti dati» quando ad esempio si vuole affermare qualcosa sulla sua continuità.[4]

Una piena consapevolezza della novità teorica insita nella propria definizione di continuità, rispetto alla tradizione euleriana e lagrangiana, si trova in Cauchy in uno scritto molto più tardo, intitolato La *Mémoire sur les fonctions continues ou discontinues*, in cui si legge:

> Nelle opere di Euler e di Lagrange, una funzione è chiamata *continua* o *discontinua*, secondoché i diversi valori di essa, corrispondenti a diversi valori della variabile, sono o non sono soggetti a una medesima legge, sono o non sono forniti da una sola e medesima equazione. È in questi termini che la continuità delle funzioni si trova definita da questi illustri geometri, allorché dicevano che le funzioni arbitrarie, introdotte dall'integrazione delle equazioni alle derivate parziali, possono essere funzioni continue o discontinue.
>
> Tuttavia, siffatta definizione è lontana dall'offrire una precisione matematica, poiché, se i diversi valori di una funzione corrispondenti ai diversi valori di una variabile dipendono da due o più equazioni distinte, nulla impedirà di diminuire il numero di queste equazioni, e anche di sostituirvi un'unica equazione la cui decomposizione fornirebbe tutte le altre. Di più: le leggi analitiche, alle quali le funzioni possono essere assoggettate, si trovano generalmente espresse da formule algebriche o trascendenti, e può accadere che diverse formule rappresentino, per certi valori di una variabile x, la stessa funzione, e per altri valori di x, delle funzioni differenti. Quindi, se si considera la definizione di Euler e di Lagrange come applicabile ad ogni specie di funzioni, un semplice cambiamento di notazione basterà sovente per trasformare una funzione continua in funzione discontinua, e reciprocamente. Così, per esempio, supposto x reale, una funzione che si riducesse ora a $+x$ ora a $-x$, secondoché x fosse positiva o negativa, dovrebbe per questo motivo annoverarsi tra le funzioni discontinue, eppure la stessa funzione potrà riguardarsi come continua quando sia rappresentata dall'integrale definito
>
> $$\frac{2}{\pi}\int_0^\infty \frac{x^2}{t^2+x^2}\,dt$$
>
> Dunque, il carattere di continuità nelle funzioni, considerato dal punto di vista dal quale dapprima si misero i geometri, è un carattere vago e indeterminato. L'indeterminazione, però, cesserà se alla definizione di Euler si sostituisce quella da me data nel secondo capitolo dell'*Analyse algébrique* [Cauchy, 1844, p. 116].

L'opera analitica di Cauchy fu importante anche perché in essa troviamo una robusta sistemazione teorica di alcune nozioni geometriche introdotte

4 La definizione di continuità formulata da Cauchy era stata data qualche anno prima da B. Bolzano (1781-1848), che l'aveva pubblicata nel 1817. Le ricerche di Bolzano passarono però inosservate per oltre cinquant'anni. Anche la sua scoperta che esistono funzioni, definite su un intervallo, continue ma non derivabili in alcun punto del dominio, venne trascurata dai matematici suoi contemporanei. L'esempio di una funzione di questo tipo, fornito da Weierstrass nelle sue lezioni universitarie del 1861 e pubblicato poi in un lavoro presentato all'Accademia di Berlino nel 1872, fu generalmente accolto come il primo del genere.

a partire dalla fine del XVII secolo in termini intuitivi, cui successivamente erano state applicate le tecniche analitiche. Le nozioni di «lunghezza di una curva», di «area delimitata da una curva», di «volume delimitato da una superficie», erano state accettate nel modo in cui venivano intese intuitivamente e veniva considerato uno dei massimi risultati del calcolo infinitesimale il fatto che queste quantità potessero essere calcolate mediante gli integrali. Cauchy, invece, *definì* queste grandezze geometriche mediante gli integrali che erano stati formulati per calcolarle.

Cauchy dedicava largo spazio allo studio delle serie: queste erano uno strumento essenziale nell'analisi così come si era andata costituendo nel Settecento e con Lagrange erano state assunte a fondamento del calcolo infinitesimale. Nel sesto capitolo del *Cours d'analyse*, dopo aver dimostrato una condizione necessaria e sufficiente per la convergenza di una serie (oggi detta «condizione di Cauchy»), Cauchy enunciava un teorema all'apparenza assai convincente: data una serie di funzioni continue in un intorno di un certo valore particolare x per il quale la serie è convergente, anche la somma della serie è una funzione continua nell'intorno di quel valore particolare.

Cauchy non fa il nome di Fourier in questa circostanza, ma è chiaro che qui egli ne mette in discussione l'intero approccio: essendo infatti le funzioni seno e coseno, che compaiono nelle serie di Fourier, funzioni continue, ne segue che, secondo questo teorema, le serie di Fourier associate a funzioni discontinue non sono convergenti alle funzioni date.

> L'errore nella «dimostrazione» di Cauchy si deve a diversi tipi di confusioni. Anzitutto, l'uso sistematico e sbrigativo di «infinitesimi» nella dimostrazione gli impedisce di cogliere le diverse dipendenze funzionali. [...] In secondo luogo, la sua concezione geometrica, sostanzialmente intuitiva, lo porta a pensare che, al crescere di n, i grafici delle [somme parziali] $y = s_n(x)$ si avvicinino sempre più a quello (supposto esistente) della somma $f(x)$. [...] Un'ultima confusione riguarda le operazioni con doppi limiti [Bottazzini, 1981, p. 106].

Il primo ad avanzare delle riserve sul teorema di Cauchy fu Abel. In una nota che compare nel suo articolo del 1826 sulla convergenza della serie binomiale egli osserva: «Mi sembra che questo teorema ammetta delle eccezioni. Per esempio la somma della serie

$$\sin(x) - \frac{1}{2}\sin(2x) + \frac{1}{3}\sin(3x) + \dots \tag{18}$$

è discontinua per tutti i valori $(2m+1)\pi$ di x, dove m è un numero intero. Ci sono, come si sa, diverse serie di questo tipo».

Quella citata da Abel era una delle serie proposte da Fourier come esempi dei suoi sviluppi in serie trigonometriche. È dopo la lettura del *Cours* che le serie di Fourier si rivelano agli occhi di Abel come effettivi controesempi del teorema di Cauchy, proprio sulla base della definizione di continuità che dà il matematico francese. Come Cauchy, anche Abel si sofferma lungamente sulle questioni relative alle serie; opponendosi a Lagrange, egli sostiene che in generale non è lecito, per derivare una funzione data mediante una serie, derivare termine a termine la serie, pensando che in questo modo si è ottenuta proprio la derivata della funzione rappresentata dalla serie di partenza:

questo, si badi, è quanto fa ripetutamente e *senza problemi* anche Fourier quando determina i coefficienti delle sue serie.[5]

Cauchy era consapevole della novità teorica delle sue idee sui fondamenti dell'analisi, radicalmente diverse da quelle lagrangiane. Nelle pagine introduttive del *Cours*, Cauchy denunciava apertamente l'insufficienza del «metodo delle serie di Taylor» per fondare il calcolo, contrariamente a quanto aveva affermato Lagrange nella *Théorie*: «Malgrado tutto il rispetto che si deve a una grande autorità — osservava a questo proposito Cauchy — la maggior parte dei geometri oggi concorda nel riconoscere l'incertezza dei risultati ai quali si può essere portati usando le serie divergenti». In particolare, in diversi casi la serie di Taylor «sembrava fornire lo sviluppo di una funzione in serie convergente, sebbene la somma della serie differisca in maniera essenziale dalla funzione proposta [Cauchy, 1823, p. 2]». Era questo il sorprendente risultato cui egli era giunto considerando l'esempio della funzione

$$f(x) = \exp(-1/x^2) \quad \text{se } x \neq 0 \qquad f(0) = 0 \tag{19}$$

le cui derivate, per $x = 0$, sono tutte nulle. Così, mentre per Lagrange la serie di Taylor era la *premessa* della teoria delle funzioni derivate, per Cauchy lo studio della possibilità di sviluppare una funzione in serie di potenze costituisce la conclusione delle sue lezioni sul calcolo infinitesimale.[6]

In questo modo, Cauchy non solo metteva definitivamente in luce la debolezza intrinseca della costruzione di Lagrange, ma inaugurava una delle questioni più spinose dell'analisi dell'Ottocento, quella di stabilire se e come si potevano rappresentare le funzioni in serie convergenti alla funzione data. La ricerca e la determinazione di criteri di convergenza per le serie riflette un radicale mutamento nel modo di intendere l'analisi rispetto alla tradizione settecentesca: Cauchy si assicura della convergenza delle serie sotto opportune ipotesi per il termine generale *senza conoscere effettivamente* quale sia il valore somma. Le affermazioni di natura esistenziale sono diventate oggi abituali in matematica e costituiscono anzi la struttura portante di numerose teorie (e solo chi aderisce ad un punto di vista strettamente costruttivista o neointuizionista è portato a metterne in discussione la validità). In questo senso si trova qui in Cauchy una profonda novità teorica, destinata a improntare di sé lo sviluppo futuro della matematica (a Cauchy si devono, fra l'altro, teoremi di esistenza e unicità della soluzione di equazioni differenziali).

Si trattava, ora, per quanto riguarda le serie di Fourier (argomento all'epoca di grande interesse teorico), di stabilire teoremi di convergenza, senza essere costretti caso per caso, come era stato nell'esemplificazione prodotta dallo stesso Fourier, a calcolare concretamente il limite per ogni serie trigonometrica di cui si voglia stabilire la convergenza.

1.8 FUNZIONE DI DIRICHLET

L'opera di Fourier aveva rivelato che una vasta classe di funzioni può essere rappresentata mediante serie trigonometriche. Rimaneva però aperto

5 La natura della convergenza della serie di Fourier ricevette ulteriore attenzione dopo l'introduzione da parte di Weierstrass, Stokes (1819-1903) e Seidel (1821-1896) del concetto di «convergenza uniforme» [Bottazzini, 1981, p. 158].

6 Cauchy aveva presentato il suo controesempio del «teorema di Taylor» all'Accademia delle Scienze di Parigi nel 1822 e poi in una nota apparsa lo stessa anno nel *Bullettin de la Société Philomatique*.

il problema di determinare condizioni sufficienti affinché una funzione sia sviluppabile in serie di Fourier. Gli sforzi compiuti da Cauchy e da Poisson non diedero frutti.

La questione della convergenza delle serie di Fourier fu affrontata da Dirichlet nel lavoro *Sur la convergence des séries trigonometriques* [1829]; qui Dirichlet diede il primo insieme di condizioni sufficienti («condizioni di Dirichlet») affinché la serie di Fourier di una funzione f periodica e continua converga puntualmente ad f. La dimostrazione data da Dirichlet è un affinamento di quella abbozzata da Fourier nelle sezioni conclusive della *Théorie analytique de la chaleur*. Come esempio di funzione che non soddisfa questa condizione egli dà la celebre «funzione di Dirichlet», una funzione di x che vale c (costante) quando x è razionale e d (costante $\neq c$) quando x è irrazionale.

> La funzione così definita ha dei valori finiti e determinati per ogni valore di x e nondimeno non sarebbe possibile sostituirla nella serie, dal momento che i diversi integrali che entrano in questa serie perderebbero in questo caso ogni significato [Dirichlet, 1829, p. 132].

La restrizione sulle discontinuità della funzione è legata alla nozione di integrale di Cauchy, allora la sola a disposizione. Cauchy aveva definito l'integrale per funzioni continue su un intervallo o discontinue in un numero finito di punti.

Dirichlet sembra essere dell'opinione che una funzione continua sia *sempre* rappresentabile in serie di Fourier. Quando più di vent'anni dopo, nel 1853, Gauss ritorna sull'argomento e gli scrive, ipotizzando la possibilità di estendere la dimostrazione della convergenza delle serie trigonometriche a tutte le funzioni continue, Dirichlet non esita a rispondere: «Dopo un esame più ravvicinato della cosa, trovo completamente confermata la Sua supposizione, se si vuole in qualche modo prescindere da certi casi del tutto singolari».

La lettera continua con una traccia di come si potrebbe condurre in porto la dimostrazione. Che l'ottimismo di Dirichlet sulla portata del suo teorema fosse mal posto sarà chiaro solo nel 1876, quando Du Bois-Reymond (1831-1889) renderà noto il controesempio di una funzione continua la cui serie di Fourier non converge puntualmente alla funzione data; in questa circostanza Du Bois-Reymond confermava di aver appreso da una conversazione con Weierstrass «che Dirichlet sembrò non aver mai perso la convinzione della validità del suo teorema».[7]

Dirichlet stesso, tuttavia, si rendeva conto che l'esame dei casi esclusi dalle sue condizioni di convergenza coinvolgeva questioni fra le più delicate dell'analisi, quali la continuità e l'integrabilità di una funzione, e chiudeva la sua memoria con l'osservazione che «la cosa, per essere fatta con tutta la chiarezza desiderabile, esige qualche dettaglio legato ai principi fondamentali dell'analisi [Dirichlet, 1829, p. 132]».

Dirichlet pubblicherà diversi anni più tardi, nel 1837, un secondo articolo sull'argomento, in cui, fra l'altro, riportava la definizione di alcuni concetti fondamentali di cui faceva uso: quello di integrale (nel senso di Cauchy) e quello di continuità di una funzione in un intervallo. La notorietà acquisita da questo lavoro risiede probabilmente in quest'ultima definizione, come

7 Il controesempio è riportato in [Bottazzini, 1981, p. 168]. Esempi più semplici di funzioni continue ma non rappresentabili in serie di Fourier furono dati da Schwartz (vedi [Sachse, 1880, p. 272] e [Lebesgue, 1906, p. 85]).

anche Kronecker lascia intendere, chiamandola semplicemente «la memoria sulla continuità», mentre l'effettivo argomento, desunto dal titolo, era la «rappresentazione di funzioni del tutto arbitrarie mediante serie di seni e coseni». Ecco la definizione in questione:

> Si pensi di indicare con a e b due valori fissati e con x una grandezza variabile, che possa assumere tutti i valori compresi fra a e b. Ora, ad ogni x, corrisponda un unico y finito e tale che, mentre x percorre con continuità l'intervallo da a a b, $y = f(x)$ vari in maniera del tutto simile; allora y si dice *funzione continua* [corsivo nostro] di x in quest'intervallo. Per ciò, non è affatto necessario che y sia dipendente da x secondo la stessa legge nell'intero intervallo; neppure occorre pensare a una dipendenza esprimibile attraverso operazioni matematiche.
>
> Rappresentata geometricamente, cioè pensate la x e la y come ascisse e ordinate, una funzione continua appare come una curva connessa (*zusammenhängende Kurve*) di cui ad ogni valore dell'ascissa compreso fra a e b corrisponde un solo punto.
>
> Questa definizione non impone un'unica legge per le varie parti della curva; questa si può pensare composta delle più diverse parti o disegnata del tutto arbitrariamente. Ne segue che una simile funzione si deve considerare completamente determinata su un intervallo quando o è data graficamente per l'intero intervallo oppure, matematicamente, quando è sottoposta per le varie parti dell'intervallo a leggi valide in esse. Anche quando si sia determinata una funzione per una parte dell'intervallo, rimane del tutto arbitraria la maniera in cui prolungare la funzione nella parte restante dell'intervallo [Dirichlet, 1837, p. 135].

Commentando questa definizione, Hankel scriveva nel 1870 che i risultati di Fourier, che avevano ampliato la classe delle funzioni ammissibili in analisi alle funzioni discontinue, avevano definitivamente rivelato insostenibile l'ipotesi tacita, ma decisiva, che le proprietà delle funzioni analitiche si potessero comunque estendere a tutte le funzioni: perciò il vecchio concetto di funzione, che richiedeva che una qualunque funzione fosse rappresentabile analiticamente, si rivelava inadeguato. «Reciso questo nodo», dice Hankel, la via era spianata per una definizione di funzione come la seguente:

> Si dice che y è funzione di x se ad ogni valore della grandezza variabile x all'interno di un certo intervallo corrisponde un determinato valore di y, senza riguardo al fatto che su tutto l'intervallo y dipenda o no da x secondo la stessa legge, e che la dipendenza sia o no esprimibile da operazioni matematiche [Hankel, 1870, p. 63].

È questa la forma in cui Hankel presenta la definizione di Dirichlet, andando ben oltre, dunque, le intenzioni e lo scritto dello stesso Dirichlet. Scrive infatti Hankel:

> Questa definizione puramente nominale, cui nel seguito assocerò il nome di Dirichlet, poiché essa sta a fondamento del suo lavoro sulle serie di Fourier, che hanno mostrato senza ombra di dubbio l'indifendibilità del precedente concetto, non basta affatto alle esigenze dell'analisi, poiché funzioni di questo tipo non possiedono proprietà generali e perciò vengono meno tutte le relazioni fra i valori delle funzioni corrispondenti ai diversi valori dell'argomento [Hankel, 1870, p. 64].

Così, dice Hankel, c'è una grande confusione su che cosa si debba intendere come funzione, come si vede anche da un rapido sguardo ai migliori manuali: c'è chi la definisce come Euler, chi come Dirichlet, chi dice che una dipendenza funzionale è data secondo una «legge» e chi non dà alcuna

definizione; «tutti tuttavia traggono dai loro concetti conseguenze che non sono affatto contenute in essi [Hankel, 1870, p. 65]».

C'era dunque bisogno, continua Hankel, di dare una definizione di funzione che consentisse di stabilire una teoria e un calcolo con le funzioni coerente con la definizione data, non troppo generale da diventare «puramente nominale e priva di contenuto reale» e tuttavia in grado di individuare una vasta classe di oggetti matematicamente interessanti e significativi. Per Hankel la classe di funzioni «che risponde nella maniera più completa possibile ai bisogni dell'analisi» è quella delle funzioni meromorfe di una variabile complessa, la cui teoria aveva trovato una solida sistemazione concettuale con l'opera di Cauchy, Riemann e Weierstrass.

Anche K. Weierstrass (1815-1897), nelle sue lezioni introduttive alla teoria delle funzioni analitiche [1886], andava sostenendo che la definizione di Dirichlet, a causa della sua grande generalità, permetteva in pratica di «fare poche cose», e se comunque veniva utilizzata era perché si trasportavano «tacitamente le proprietà che possiedono tutte le funzioni considerate a queste funzioni generali».

Le interpretazioni di Hankel e Weierstrass della definizione di funzione *continua* (ma come si è visto questa precisazione è subito scomparsa nei commenti dei due) data da Dirichlet sono state fatte proprie dai moderni storici della matematica: così E. T. Bell vi ha visto la prima formulazione del moderno concetto di applicazioni fra insiemi e ha scritto (1945) che «la definizione di Dirichlet di una funzione (a valori numerici) di una variabile reale come una tavola o corrispondenza, o correlazione, fra due insiemi di numeri, lascia intendere una teoria dell'equivalenza degli insiemi di punti», dando con ciò prova di una notevole libertà interpretativa. C. B. Boyer, da parte sua, dà questa versione dei fatti:

> Lejeune-Dirichlet [...] nel 1837 propose una definizione molto ampia di funzione: se una variabile y ha una relazione con una variabile x tale che, ogniqualvolta venga assegnato un valore numerico alla x, esista una regola in base alla quale viene determinato un valore univoco di y, si dice che y è una funzione della variabile indipendente x.
>
> Questa definizione presenta un'affinità con l'idea moderna di corrispondenza fra due insiemi di numeri, anche se a quel tempo i concetti di «insieme» e di «numero reale» non erano ancora entrati stabilmente nel campo della matematica. Per sottolineare la natura assolutamente arbitraria della regola di corrispondenza, Dirichlet propose una funzione molto «irregolare»: quando x è razionale, sia $y = c$, e quando x è irrazionale, sia $y = d \neq c$ [Boyer, 1968, p. 635].

Più sbrigativamente, Bourbaki si limita a osservare che «è noto che Dirichlet, precisando le idee di Fourier, definisce la nozione generale di funzione come noi l'intendiamo oggi [Bourbaki, 1960, p. 237]». A questo proposito osserva però Bottazzini [1981]:

> Tutto ciò è molto suggestivo, ma purtroppo non ha nulla a che fare con la storia reale: né Dirichlet propose quella definizione né tantomeno diede l'esempio di funzione «patologica» citato per illustrare l'arbitrarietà della regola di corrispondenza. Quello che sembra essere loro sfuggito in primo luogo è l'aggettivo «continua» che accompagna il termine funzione, senza il quale appunto si finisce per non comprendere nulla delle idee di Dirichlet. Egli è interessato a definire esplicitamente ciò che Fourier intendeva per funzione arbitraria e si collega alla lunga discussione originata dalla definizione euleriana di funzione continua e dalla nuova definizione di continuità (alla Cauchy).

> Ciò che Dirichlet vuole chiarire è semplicemente che una funzione continua si può dare o arbitrariamente con un grafico (e qui la sua idea di funzione continua mostra di essere largamente intuitiva) oppure con una formula matematica, non necessariamente la stessa in ogni parte dell'intervallo [Bottazzini, 1981, p. 156].

L'idea di fondo di Dirichlet è che ogni funzione continua, per quanto arbitrariamente data, sia sviluppabile in serie di Fourier. Quanto Dirichlet fosse distante da quello che oggi viene generalmente chiamato «il concetto di Dirichlet di funzione» emerge anche da ciò che egli scrive nella stessa memoria del 1837; quando si tratta di discutere il valore della somma della serie di Fourier nei punti di discontinuità della funzione, egli asserisce:

> La curva la cui ascissa è β e la cui ordinata è $f(\beta)$ consiste di più pezzi, la cui connessione è interrotta nei punti dell'asse delle ascisse che corrispondono a quei particolari valori di β, e per ognuna di tali ascisse corrispondono di fatto due ordinate, di cui l'una appartiene alla porzione di curva che termina in quel punto e l'altra alla porzione che vi comincia. Nel seguito sarà necessario distinguere questi *due valori di* $f(\beta)$ [corsivo nostro] e li indicheremo con $f(\beta - 0)$ e $f(\beta + 0)$ [Dirichlet, 1837, p. 156].

Dopo aver provato che la serie di Fourier di una funzione f, periodica e continua a tratti, converge in ogni punto $x \in \mathbb{R}$ in cui sono soddisfatte le «condizioni di Dirichlet», con somma

$$s(x) = \frac{1}{2}[f(x+0) + f(x-0)] \tag{20}$$

(secondo la notazione introdotta), Dirichlet così commenta:

> Dove si presenta una soluzione di continuità e dunque la funzione *ha propriamente due valori* [corsivo nostro] la somma della serie, che per propria natura assume per ogni x un valore univoco, rappresenta la semisomma di questi valori [Dirichlet, 1837, p. 159].

Sono al più di questo tipo le funzioni «del tutto arbitrarie» che compaiono già nel titolo delle due memorie e che Dirichlet considera. L'origine dell'attribuzione a Dirichlet del «concetto di Dirichlet» di funzione è verosimilmente da attribuire al suo esempio di funzione «patologica», anche se Dirichlet stesso non sembrava attribuire grande interesse a questa funzione, che non rientrava neppure fra quelle «completamente arbitrarie», per le quali aveva senso parlare di integrale e rappresentazione in serie di Fourier. Il quadro concettuale cambierà solo con la fondazione del calcolo infinitesimale sul concetto di «numero reale», che sta al centro del programma di «aritmetizzazione dell'analisi».[8]

1.9 ARITMETIZZAZIONE DELL'ANALISI

M. Kline [1972] introduce così la sua ricostruzione storica della nascita della teoria dei numeri reali:

8 È interessante osservare che la stessa funzione di Dirichlet può essere definita analiticamente, poiché essa è rappresentabile mediante l'espressione

$$f(x) = \lim_{m\to+\infty} \lim_{n\to+\infty} \cos(m!\pi x)^n = \begin{cases} 1 & \text{se } x \in \mathbb{Q} \\ 0 & \text{altrimenti} \end{cases}$$

come G. Peano e R. Baire hanno dimostrato.

> Uno dei fatti più sorprendenti della storia della matematica è che la fondazione logica del sistema dei numeri reali sia stata edificata soltanto alla fine del XIX secolo. Fino ad allora non solo non erano state stabilite nemmeno le proprietà più semplici dei numeri reali, ma questi numeri non erano neppure stati definiti. Neanche la fondazione logica dei numeri complessi [compiuta da W. R. Hamilton nel 1837, con il lavoro *Conjugate Functions and on Algebra as the Science of Pure Time*] aveva lunga vita e anch'essa presupponeva il sistema dei numeri reali [Kline, 1972, p. 1143].

Se pensiamo all'esteso sviluppo dell'analisi e dell'algebra, che di fatto utilizzavano entrambe i numeri reali, la mancata considerazione della struttura precisa e delle proprietà dei numeri reali indica quanto poco «logicamente» progredisca la matematica nel suo sviluppo storico. La comprensione intuitiva di questi numeri sembrava adeguata e i matematici si accontentavano di operare su questa base.

Nella seconda metà del secolo scorso, di fronte alla straordinaria ricchezza di risultati, i matematici più attenti alle questioni metodologiche cominciarono però a rendersi conto che anche le nozioni apparentemente più sicure e i teoremi fondamentali dell'analisi mancavano di una base adeguatamente rigorosa. Così era per esempio per il criterio di convergenza delle serie dato da Cauchy, come anche per il teorema secondo cui ogni successione monotona e limitata ammette limite, un teorema che «può essere considerato in un certo senso come base sufficiente del calcolo differenziale», come si esprime Dedekind (1831-1916) nella prefazione a *Continuità e numeri irrazionali* [1872]. La stessa nozione di continuità appariva incerta e lo studio delle serie trigonometriche lontano dal potersi pensare concluso.[9]

La convinzione che lentamente si venne affermando, non senza polemiche e opposizioni, fu che il rigore desiderato si poteva raggiungere costruendo e strutturando in maniera precisa il sistema dei numeri reali, partendo dai numeri razionali e, a monte, dai naturali; di conseguenza, il concetto di numero naturale e l'aritmetica dei naturali venivano considerati il fondamento su cui costruire l'edificio dell'analisi.

Che cosa si debba intendere quando si afferma che il concetto di numero naturale e l'aritmetica dei numeri naturali costituiscono il fondamento dell'analisi lo esprime con l'abituale chiarezza Dedekind [1888], scrivendo che questa concezione

> [...] fa apparire come ovvio che ogni teorema di analisi, per quanto remoto, si possa enunciare come teorema sui numeri naturali. È questa un'asserzione che ho sentito fare parecchie volte anche da Dirichlet. Io, però, non considero in nessun modo come cosa meritevole, e anche Dirichlet non ha mai avuto l'intenzione di farlo, l'intraprendere effettivamente una tale faticosa opera di riduzione e il non voler ammettere ed usare altri numeri all'infuori di quelli naturali [Dedekind, 1888, p. 3].

Il punto è un altro: qui si tratta solo di assicurare, almeno in linea di principio, l'esistenza della possibilità di tale riduzione e con ciò, per Dedekind, l'esistenza di un fondamento sicuro per l'intera analisi. L'introduzione di nuovi concetti ha invece un'importanza fondamentale nello sviluppo delle teorie matematiche; di ciò Dedekind è assolutamente convinto:

9 Bolzano, sin dal 1817, si era reso così pienamente cosciente dell'esigenza di rigore nel campo dell'analisi che Klein lo definì il «padre dell'aritmetizzazione»; le ricerche di Bolzano passarono però inosservate per oltre cinquant'anni.

> Si constata invece — egli scrive — che i progressi più grandi e fecondi nella matematica e nelle altre scienze sono dovuti prevalentemente alla formazione e all'introduzione di nuovi concetti, perché ciò è reso necessario per la frequente riapparizione di fatti complicati, che vengono a stento dominati dai vecchi concetti [Dedekind, 1888, p. 5].

I mutamenti negli standard di rigore, così come la tendenza alla generalizzazione e l'introduzione di nuovi concetti, trovano dunque motivazione nei problemi affrontati e si misurano sul terreno concreto della pratica matematica.

Alle ricerche sul continuo e su una sistemazione rigorosamente aritmetica della continuità si dedicarono molti importanti matematici fra il XIX e il XX secolo; in particolare ricordiamo C. Méray (1835-1911), K. Weierstrass (1815-1897), H. E. Heine (1821-1881), G. Cantor (1845-1918) e R. Dedekind (1831-1916). Costoro, in un certo senso, rappresentano il culmine di oltre due secoli di ricerche sulla natura della funzione e del numero, cominciate verso la metà del Seicento con il «metodo delle flussioni» di Newton. Le teorie dei numeri reali da essi elaborate sono per molti aspetti simili e ci limiteremo perciò a dare alcune indicazioni sulle teorie di Weierstrass e Dedekind.[10]

Weierstrass [1872] si rese conto che, per evitare di cadere in una *petitio principii*, era necessario dare una definizione di numero reale che fosse indipendente dal concetto di limite, dal momento che quest'ultimo sino ad allora aveva presupposto il primo. Weierstrass risolse la questione dell'*esistenza* del limite di una successione «di Cauchy» a valori razionali facendo della successione stessa il numero reale ovvero il limite. Nella teoria di Weierstrass, sostanzialmente, le successioni «di Cauchy» che non convergevano a nessun numero razionale venivano considerate come *definienti* numeri irrazionali.

Un metodo completamente diverso di affrontare lo stesso problema (metodo che oggi è quello più conosciuto) fu presentato nello stesso anno da Dedekind in un libro diventato famoso, *La continuità e i numeri irrazionali* [1872]. Dedekind scrive nella prefazione al suo saggio:

> Spesso si dice che il calcolo differenziale si occupa di grandezze continue, eppure non si dà mai una definizione di questa continuità. Le trattazioni più rigorose che si hanno del calcolo differenziale non basano le loro dimostrazioni sulla continuità, ma fanno invece appello, più o meno coscientemente, a rappresentazioni geometriche o si servono di teoremi che a loro volta non furono mai rigorosamente dimostrati con mezzi *puramente aritmetici* [corsivo nostro] [Dedekind, 1872, p. 2].

Rifiutando il ricorso all'evidenza geometrica, Dedekind sottolinea la necessità di «scoprire negli elementi dell'aritmetica la vera origine» dei teoremi del calcolo differenziale, «acquistando con ciò nello stesso tempo una definizione effettiva della continuità».

Come Weierstrass, anche Dedekind, nella costruzione del sistema dei numeri reali, presuppone la teoria dei numeri razionali. Per affrontare i numeri reali, Dedekind si chiede anzitutto che cosa si intenda per continuità «geometrica». La via da lui seguita si basa sulla proprietà di ordinamento della retta, mediante la quale si arriva a formulare «una proprietà caratteristica

10 Va però sottolineato che, nella teoria dei numeri reali di Cantor, la corrispondenza biunivoca fra punti della retta e numeri reali è assicurata mediante un assioma di continuità analogo a quello di Dedekind, ma non equivalente: il postulato di continuità della retta di Dedekind implica quello di Archimede, che non è derivabile dall'assioma di continuità di Cantor [Carruccio, 1972, p. 230].

e precisa della continuità, che possa servire di base a deduzioni vere e proprie». Infatti ogni punto della retta divide la retta in due parti, tali che ogni punto di una di esse precede, nell'ordine naturale, ogni punto dell'altra.

> Ora — scrive Dedekind — io vedo l'essenza della continuità nell'inversione di questa proprietà, e cioè nel principio seguente: «Se una ripartizione di tutti i punti della retta in due classi è di tale natura che ogni punto di una delle due classi sta a sinistra di ogni punto dell'altra, allora esiste uno e un sol punto dal quale questa ripartizione di tutti i punti in due classi, o questa decomposizione della retta in due parti, è prodotta [Dedekind, 1872, p. 14]».

È un principio tanto evidente, dice Dedekind, che la maggior parte delle persone lo riterrà assai banale e «proverà una grande disillusione nell'apprendere che è questa banalità che deve svelare il segreto della continuità». Che tale principio sia altamente evidente è però al più elemento di soddisfazione, poiché esso «non è che un assioma, ed è sotto forma di questo assioma che noi pensiamo la continuità della retta».

Definita di Dedekind in maniera assiomatica la continuità della retta, si è trovata anche la via da percorrere per giungere ad una caratterizzazione aritmetica di essa. Consideriamo, con Dedekind, l'insieme $\mathbb{Q}$ dei razionali, denso e ordinato. Diremo che una coppia $(\mathcal{A}, \mathcal{B})$ di sottoinsiemi di $\mathbb{Q}$ è una *sezione* in $\mathbb{Q}$ se valgono le tre seguenti proprietà:

(*i*) $(\mathcal{A}, \mathcal{B})$ è una partizione di $\mathbb{Q}$;

(*ii*) ogni elemento di $\mathcal{A}$ precede ogni elemento di $\mathcal{B}$;

(*iii*) $\mathcal{B}$ non ha un elemento minimo.

Se $\mathcal{A}$ ha massimo $q \in \mathbb{Q}$, diciamo che la sezione $(\mathcal{A}, \mathcal{B})$ è *generata* dal numero razionale q.

Si verifica immediatamente che c'è una corrispondenza biunivoca fra $\mathbb{Q}$ e l'insieme delle sezioni generate dai numeri razionali. Fatto questo, Dedekind dà un esempio di un'infinità di sezioni che non sono generate da numeri razionali.

> È facile provare l'esistenza di infinite sezioni non prodotte da nessun numero razionale: un esempio immediato è dato dalla sezione $(\mathcal{A}, \mathcal{B})$ tale che in $\mathcal{B}$ stiano tutti i numeri razionali positivi il cui quadrato è maggiore di a, con a intero positivo non quadrato perfetto, e nella classe $\mathcal{A}$ stiano i rimanenti numeri razionali [Dedekind, 1872, p. 23].

Ciascuna di queste sezioni si dice *numero irrazionale*. L'insieme delle coppie $(\mathcal{A}, \mathcal{B})$ è detto *insieme dei numeri reali* $\mathbb{R}$. La costruzione è completata dando ad $\mathbb{R}$ la struttura di campo ordinato.[11]

11 Il linguaggio usato da Dedekind per introdurre i numeri irrazionali non è del tutto rigoroso. Dedekind scrive che

> [...] ogni volta che è data una sezione $(\mathcal{A}, \mathcal{B})$ che non sia prodotta da nessun numero razionale noi creiamo un nuovo numero, un numero irrazionale α, che noi consideriamo completamente definito da questa sezione; noi diremo che il numero α corrisponde a questa sezione e che esso la produce [Dedekind, 1872].

Dedekind introduce l'irrazionale α come «corrispondente» alla sezione e «prodotto» dalla sezione, ma non è proprio chiaro che cosa sia α. In effetti, il numero irrazionale non è nulla di più che la sezione. H. Weber fece questa obiezione a Dedekind in una lettera del 1888, ma Dedekind rispose che il numero irrazionale α non è la sezione, ma qualche cosa di distinto che corrisponde alla sezione e che provoca la sezione. Analogamente, sebbene generino delle sezioni, i numeri razionali non sono la stessa cosa delle sezioni. Dedekind dice che noi possediamo il potere mentale di «creare» tali concetti.

Dedekind dimostra poi come, per mezzo dei numeri reali, si possa dare una dimostrazione rigorosa di un fondamentale teorema del calcolo infinitesimale, quello dell'esistenza del limite per una successione monotona e limitata. Dedekind sottolineava il fatto che, in tal modo, era possibile dimostrare rigorosamente i teoremi fondamentali dell'analisi senza fare ricorso a considerazioni geometriche intuitive:

> Era stata la geometria ad indicare la via per giungere ad una adeguata definizione del concetto di continuità, ma alla fine essa veniva esclusa dalla formale definizione aritmetica di tale concetto [Boyer, 1968, p. 644].

Le «sezioni» di Dedekind nel sistema dei numeri razionali, oppure una costruzione equivalente del concetto di numero reale, venivano così a costituire la spina dorsale dell'analisi.

La questione della fondazione del calcolo infinitesimale sul concetto di «numero reale», che è al centro del programma di aritmetizzazione dell'analisi, si colloca nel contesto generale dell'esigenza di rigore nella costruzione dei fondamenti logici dell'analisi e quindi nella definizione rigorosa dei concetti con cui si opera; così non stupisce che Kronecker (1823-1891) affermi che

> [...] i matematici sono un po' supponenti (*hochmütig*) nell'uso del concetto di funzione. Anche Riemann, generalmente molto esatto, non è irreprensibile sotto questo aspetto. Se una funzione diminuisce e poi cresce o viceversa, Riemann dice dovervi essere un minimo o un massimo (vedi la dimostrazione del cosiddetto «principio di Dirichlet»), mentre si dovrebbe restringere la conclusione alla sfera delle funzioni per così dire ragionevoli [Kronecker, 1887, p. 24].

L'assunto che qui è implicito è l'esigenza di svincolare definitivamente la fondazione logica dei concetti dell'analisi dal riferimento a considerazioni geometriche intuitive; in particolare, il concetto di funzione da quello intuitivo di curva piana. Acquistano così particolare interesse le funzioni «patologiche» sulle quali aveva richiamato l'attenzione Riemann (1826-1866).

Lo studio ravvicinato di funzioni dal comportamento eccezionale o «patologico» ha così lo scopo di chiarire e delimitare la portata dei concetti, come quello di funzione analitica e di singolarità di una funzione, quello di minimo e di estremo inferiore (nella critica che fa Kronecker all'uso riemanniano del «principio di Dirichlet»), o quello di continuità. Tutto ciò illustra bene il costante sforzo di matematici come Dedekind e Weierstrass di enunciare le proprie definizioni e dimostrare i propri teoremi con la massima precisione e il massimo rigore possibile, con un atteggiamento e uno stile che diventeranno a poco a poco dominanti.[12]

Fra i contributi di Weierstrass al programma di aritmetizzazione dell'analisi vi fu non solo una definizione soddisfacente del concetto di numero reale, ma anche un perfezionamento della definizione di limite. La definizione di Cauchy faceva uso di espressioni come «valori successivi» o «avvicinarsi indefinitamente» o «così piccolo quanto si vuole». Per quanto suggestive e pedagogicamente convincenti, queste definizioni suggeriscono infelicemente le idee di «tempo» e di «moto» e sono prive di quella precisione che generalmente si esige nella matematica moderna. Nelle sue lezioni (1861),

12 Per una ricostruzione storica delle problematiche relative al «principio di Dirichlet» cfr. [Bottazzini, 1981, p. 242].

Weierstrass diede la moderna definizione di limite e di funzione continua in termini di ε e δ, ponendo l'accento su quella che alcuni hanno definito la «teoria statica della variabile»:

> Nella fredda e precisa definizione [di limite e continuità secondo Weierstrass] non viene fatto alcun cenno a entità che fluiscono generando grandezze di dimensione superiore, non si fa alcun ricorso a punti e linee in movimento, né si parla di quantità che divengono infinitamente piccole. Tale definizione non contiene altro che numeri reali, oltre alle operazioni di addizione e sottrazione, e alla relazione «minore di». Il linguaggio e il simbolismo non equivoci di Weierstrass e di Heine estromisero dal calcolo infinitesimale la nozione di variabilità e resero superfluo il consueto, persistente ricorso a quantità infinitesime fisse. Era veramente iniziata l'«Età del rigore», durante la quale i vecchi dispositivi euristici e intuitivi furono rimpiazzati da concetti logicamente precisi e criticamente vagliati. Le definizioni di limite di una funzione che si incontrano nei manuali correnti sono essenzialmente quelle introdotte da Weierstrass e Heine un secolo fa. Le cosiddette dimostrazioni delta-ed-epsilon fanno ora parte del bagaglio concettuale di ogni matematico [Boyer, 1968, p. 645].

La caratterizzazione delle funzioni continue in termini di «delta-ed-epsilon», che Dirichlet aveva cercato di *dimostrare*, diventa la *definizione* della continuità per Weierstrass. Da questo momento si può far datare la sostituzione dell'idea intuitiva di «tendere a» con disuguaglianze, che preludono a una nozione topologica di intorno. Si tratta di un procedimento non infrequente nella storia della matematica: dalla pratica emergono le proprietà centrali (e dunque più profonde) di classi di oggetti, ed è poi naturale porre a loro definizione proprio quelle proprietà; in questo senso si è anche parlato di «dialettica» della matematica [Cavaillès, 1938].

La costruzione del sistema dei numeri reali e la rigorizzazione dei concetti basilari del calcolo infinitesimale non si rivelarono il punto finale delle ricerche sui fondamenti. Il passo successivo nell'edificazione dei fondamenti fu costituito dalla definizione e dalla deduzione delle proprietà dei numeri razionali. Coloro che si occupavano dei numeri razionali assumevano generalmente che la natura e le proprietà dei numeri naturali fossero note e che il problema fosse quello di stabilire logicamente la teoria dei numeri interi e dei razionali. Il primo tentativo di elaborare una teoria organica dei numeri razionali fu compiuto da M. Ohm (1792-1872), nel *Versuch eines vollkommen consequenten Systems der Mathematik* (*Studio di un sistema di matematica completamente coerente*, 1822). In seguito Weierstrass, nelle lezioni tenute durante gli anni '60, derivò i numeri razionali dai numeri naturali introducendo i razionali come classi di equivalenza di coppie di interi e gli interi come classi di equivalenza di coppie di naturali. Questa idea fu riutilizzata successivamente da Peano. Weierstrass non sentì la necessità di chiarificare la logica dei naturali. Tuttavia, a partire dalle lezioni tenute nel 1859, affermò correttamente che, una volta ammessi i numeri naturali, non sono necessari ulteriori assiomi per costruire i numeri reali.

Con il senno di poi, possiamo riconoscere che il problema chiave nella costruzione del sistema dei numeri razionali era la fondazione dei numeri naturali e la dimostrazione delle loro proprietà. Fra coloro che si occuparono della teoria dei numeri naturali, alcuni ritenevano che essi fossero talmente fondamentali da renderne impossibile un'analisi logica. È questa la posizione assunta per esempio da Kronecker, che era motivato da conside-

razioni filosofiche. Anche Kronecker desiderava aritmetizzare l'analisi, cioè fondarla sulla teoria dei numeri naturali, ma riteneva che non fosse possibile andare al di là del riconoscimento della conoscenza dei numeri naturali, di cui l'uomo possiede un'intuizione fondamentale. Dedekind diede una teoria degli interi nel suo libro intitolato *Was sind und was sollen die Zahlen?* (*Che cosa sono e che cosa dovrebbero essere i numeri?*) e che, sebbene pubblicato nel 1888, risale al periodo compreso fra il 1872 e il 1878. In esso vengono usate idee provenienti dalla teoria degli insiemi che in quel momento Cantor aveva già sviluppato abbastanza a fondo e che avrebbe assunto in seguito grande importanza.

L'approccio alla teoria dei numeri naturali che meglio si adattava alle tendenze assiomatizzanti della fine dell'Ottocento era quello consistente nell'introdurli mediante un sistema di assiomi; riprendendo alcuni risultati ottenuti da Dedekind nel libro citato sopra, Peano (1858-1932) fu il primo a farlo nei suoi *Arithmetices Principia Nova Methodo Exposita* (1889).

Una volta stabilito l'approccio logico ai numeri naturali, il problema di costruire i fondamenti del sistema dei numeri reali poteva dirsi completamente risolto. È interessante osservare che i matematici che si occupavano della teoria degli irrazionali assumevano generalmente che la teoria dei numeri razionali fosse compresa con tale chiarezza da poterla prendere per data. Dopo che Hamilton ebbe basato i numeri complessi sui numeri reali e dopo che questi ultimi furono definiti mediante i numeri razionali e i razionali mediante gli interi e i naturali, venne infine creata la logica di questi ultimi. L'ordine storico fu l'esatto contrario dell'ordine logico richiesto per costruire il sistema dei numeri complessi.

Si tratta di un modo di procedere estremamente indicativo della mentalità dei protagonisti di queste ricerche: uomini concretamente impegnati a «far matematica» in modi nuovi tentano strade inesplorate, enunciano proposizioni, le saggiano, ritornano sugli enunciati, riprendono questioni apparentemente risolte per darne nuove soluzioni, con nuove prospettive. La matematica sembra qui molto più simile a una scienza sperimentale, che all'impalcatura ipotetico-deduttiva cui siamo abituati; quest'ultima appartiene alla fase di riorganizzazione logica delle teorie, più che al momento della scoperta vera e propria.

L'essenza degli approcci alla fondazione del sistema dei numeri reali finora descritti consiste nell'introdurre i numeri naturali e le loro proprietà e nel derivare da essi gli interi, i razionali e infine i numeri reali. La base logica di questo approccio consiste in un qualche insieme di asserzioni concernenti esclusivamente i numeri naturali, come ad esempio gli assiomi di Peano, mentre tutti gli altri numeri vengono costruiti. Hilbert (1862-1943) chiamava questo modo di procedere «metodo genetico». Pur riconoscendo che il metodo genetico ha un notevole valore pedagogico ed euristico, Hilbert sosteneva però che, da un punto di vista logico, era più sicuro applicare il metodo assiomatico direttamente al sistema dei numeri reali; Hilbert espose la sua introduzione assiomatica dei numeri reali nel fondamentale lavoro *Grundlangen der Geometrie* [1899].

Con la fondazione del sistema dei numeri reali, entrava stabilmente nel campo della matematica anche il concetto di «insieme». La teoria degli insiemi, più volte precisata nel corso della sua storia e raffinata sempre più nelle sue basi assiomatiche, onde evitare l'insorgere di antinomie e contraddizioni, finì ben presto per acquistare un ruolo centrale in tutta la matematica.

Con la rigorizzazione dei concetti fondamentali dell'analisi, centro del programma di aritmetizzazione, il lungo percorso del pensiero matematico verso la moderna formulazione del concetto di funzione poteva dirsi finalmente concluso. Tale formulazione rappresenta il punto di arrivo di oltre due secoli di ricerche sul calcolo infinitesimale, cominciate verso la metà del Seicento con le riflessioni newtoniane sul «metodo delle flussioni». Vedremo che, in un certo senso, fu specialmente in relazione ai tentativi volti ad estendere la nozione di funzione che ebbe origine la teoria della distribuzioni.

2 DISTRIBUZIONI

La teoria delle distribuzioni è un cardine della matematica del XX secolo e uno strumento fondamentale sia per la descrizione dei «fenomeni impulsivi» sia per lo studio delle equazioni differenziali. La teoria delle distribuzioni fu concepita fondamentalmente per due ragioni: in primo luogo, essa venne formulata per dare un assetto rigoroso ad una particolare categoria di «funzioni improprie», come la Delta di Dirac e le sue derivate, originariamente introdotte in modo euristico come oggetti «simili» alle funzioni, che possono essere nulli tranne che in un punto e avere ugualmente integrale non nullo. In secondo luogo, la teoria venne elaborata nell'ambito dello studio delle equazioni differenziali, al fine di estenderne la classe delle soluzioni ammissibili.

I fondamenti della teoria delle distribuzioni furono posti dal matematico sovietico S. L. Sobolev (1908-1989) che, nel 1936, fu il primo a definire correttamente le distribuzioni come funzionali e applicò con successo le distribuzioni allo studio del problema di Cauchy per l'equazione iperbolica. Fra il 1945 e il 1951 il matematico francese L. Schwartz (n. 1915) affrontò sistematicamente la costruzione della teoria delle distribuzioni, indicò una serie di sue applicazioni e la espose nella famosa monografia *Théorie des distributions* [1950-1951].

Il rapido sviluppo della teoria delle distribuzioni fu particolarmente stimolato dalle esigenze della fisica-matematica (meccanica, elettrodinamica, fisica quantistica). Oggigiorno la teoria delle distribuzioni è molto avanzata ed ha numerose applicazioni in matematica e in fisica. L'importanza della teoria è dovuta non soltanto all'interpretazione rigorosa che essa dà delle «funzioni improprie», ma soprattutto al fatto che essa «ha aperto una nuova area di ricerca matematica, creando la premessa per un impetuoso sviluppo in un gran numero di discipline matematiche, come la teoria delle equazioni differenziali, la teoria delle trasformate di Fourier e di Laplace, e l'analisi funzionale [Zemanian, 1965, p. 7]».

2.1 DALLE FUNZIONI ALLE DISTRIBUZIONI

In analisi matematica e in molti campi della fisica ha un ruolo centrale lo studio delle applicazioni fra spazi vettoriali finito-dimensionali, reali o complessi: noi chiameremo queste applicazioni «funzioni ordinarie». La nozione di funzione ordinaria è spesso associata, in modo più o meno esplicito, a qualche proprietà che viene richiesta alla relazione funzionale stessa. Così, per esempio, in certe applicazioni si richiede alle funzioni di essere continue; in altre è richiesta la differenziabilità una o più volte; in altre ancora è richiesta l'analiticità. Accade, però, che in alcuni casi la nozione di funzione ordinaria non sia sufficiente. Diamo alcuni semplici esempi di applicazioni fisiche e matematiche in cui accade ciò.

Si pensi, per esempio, ad una *distribuzione* di un numero finito di cariche o di masse puntiformi lungo una retta e al problema di determinare il campo elettrico o gravitazionale da essa generato. In questi casi, la densità di carica o di massa non può essere rappresentata da una funzione ordinaria, nean-

che fortemente discontinua: essa infatti dovrebbe avere la proprietà di essere ovunque nulla, ad eccezione dei punti in cui sono concentrate le cariche (o le masse), e allo stesso tempo tale che l'integrale su un intervallo contenente uno solo di questi punti dia il valore della carica (o della massa) corrispondente. Il problema si risolve se utilizziamo nuovi oggetti, che generalizzano il concetto di funzione ordinaria; tali oggetti devono essere effettivamente nuovi, in quanto ogni funzione nulla a meno di un numero finito di punti deve essere nulla a tutti gli effetti, in base alla identificazione tra funzioni uguali quasi dappertutto, cardine della teoria di Lebesgue. Nel caso specifico del problema descritto sopra tale nuovo oggetto è una *distribuzione*, combinazione lineare della cosiddetta «Delta di Dirac» in x_j ($j = 1, 2, \dots, N$): essa è precisamente la distribuzione $\sum_{j=1}^{N} c_j \delta(x - x_j)$.

Un altro esempio di «fenomeno impulsivo» è l'urto di un punto materiale P di massa m in moto rettilineo uniforme su un piano orizzontale (senza attrito) contro una parete perpendicolare alla direzione del moto. Supponendo l'urto perfettamente elastico, richiederemo, non volendo rinunciare alle leggi fondamentali della meccanica, che sia soddisfatta l'uguaglianza

$$m(v_2 - v_1) = \int_{t_1}^{t_2} F(t)\, dt \tag{21}$$

ove v_j è la velocità di P all'istante t_j ($j = 1, 2$) e $F(t)$ è la forza complessiva che agisce su P all'istante t. Denotiamo con t_0 l'istante in cui avviene l'urto ed esaminiamo le varie situazioni possibili. Se $t_1 < t_2 < t_0$ oppure se $t_0 < t_1 < t_2$, allora non ci sono difficoltà in quanto $v_1 = v_2$ e F è nulla in $[t_1, t_2]$; se invece $t_1 < t_0 < t_2$, il primo membro della (21) è $2mv_2 \neq 0$ in quanto $v_1 = -v_2 \neq 0$, mentre il secondo è nullo in quanto $F(t) = 0$ per ogni $t \neq t_0$. Anche qui il problema si risolve se utilizziamo una distribuzione. Nel caso specifico dell'urto descritto sopra tale distribuzione è multipla della Delta di Dirac in t_0: essa è la distribuzione $2mv_2\delta(t - t_0)$.

Ancora, in meccanica quantistica è spesso utile trovare una base ortonormale dello spazio delle funzioni d'onda in termini di autofunzioni di qualche operatore lineare; se lo spettro dell'operatore è continuo, e quindi la base ha cardinalità più che numerabile, non è possibile esprimere la condizione di ortonormalità in termini di funzioni ordinarie.

Per quanto riguarda le equazioni differenziali alle derivate parziali, l'opportunità di generalizzare le impostazioni «classiche» (cioè non distribuzionali) si rivela impellente nello studio di fenomeni ondosi, per i quali si impone la necessità di accettare anche funzioni discontinue. Le teorie classiche si prestano male a questo scopo: basti infatti prendere in considerazione la funzione caratteristica di $[0, +\infty[$, che denotiamo con H, e la funzione $u = 2H$, che sono trattate esattamente nello stesso modo dalla definizione classica di derivata, dato che entrambe hanno derivata nulla in $\mathbb{R} \setminus \{0\}$ e non sono derivabili in 0. Una «buona definizione» di soluzione discontinua deve invece tenere conto anche del valore del salto della funzione considerata.

Altri inconvenienti, inoltre, intervengono nella teoria classica e alcuni di essi sembrano addirittura difetti di impostazione [Gilardi, 1994]. Si pensi ad esempio alle due equazioni differenziali

$$\frac{\partial}{\partial x}\frac{\partial u}{\partial y} = 0 \qquad \frac{\partial}{\partial y}\frac{\partial u}{\partial x} = 0 \tag{22}$$

apparentemente equivalenti fra loro. Tale equivalenza si ha infatti solo rinunciando a cercarne soluzioni poco regolari: se u è di classe C^2 allora si può

usare il teorema di Schwarz e concludere che u risolve la prima equazione se e solo se risolve la seconda. Se invece si interpretano alla lettera le due equazioni differenziali e si accettano di conseguenza tutte le funzioni che rendono leciti i calcoli necessari e che risolvono le equazioni stesse, vediamo che una funzione u della forma $u(x, y) = \varphi(x)$, con φ funzione non derivabile di una variabile, è soluzione della prima equazione ma non della seconda.

Ancora, esistono funzioni derivabili in alcune direzioni e non in altre, così come esistono funzioni derivabili in tutte le direzioni ma in qualche modo «patologiche»: se non facciamo ipotesi di continuità sulle derivate, o almeno ipotesi di differenziabilità sulle funzioni considerate, le derivate direzionali possono non essere esprimibili come combinazioni lineari delle derivate parziali.

Continuiamo con l'elenco. Se una serie di funzioni converge solo puntualmente si può dire ben poco: non si può in generale integrare per serie, né si può concludere alcunché sul comportamento delle derivate. Per poter trarre conclusioni interessanti occorrono informazioni ulteriori, ad esempio legate all'uniformità della convergenza della serie data oppure della serie delle derivate.

Queste considerazioni permettono di trarre una conclusione di carattere generale: la teoria classica funziona molto bene quando si fanno ipotesi di continuità in tutti gli oggetti che intervengono (funzioni o loro derivate), mentre presenta alcune limitazioni senza tali ipotesi.

C'è un altro fatto molto importante: dato che, nella teoria di Lebesgue, due funzioni vengono identificate quando differiscono solo su un insieme di misura nulla, già la definizione di derivata non è più adatta alla situazione. Infatti, cambiando su un insieme di misura nulla una funzione derivabile, troviamo una funzione non derivabile e l'idea di cercare il «rappresentante qualificato» da scegliere sistematicamente nella definizione di derivata appare poco fruttuosa, dato che limiterebbe la nozione alle funzioni continue, mentre proprio il problema dell'urto suggerisce l'opportunità di avere un'accelerazione come derivata della velocità che «salta». Difficoltà e problemi di questo tipo possono essere risolti generalizzando opportunamente la nozione di funzione ordinaria: questa generalizzazione conduce al concetto di «distribuzione».

La teoria delle distribuzioni si occupa dello studio di funzionali lineari e continui in opportuni spazi di funzioni ed estende a tali funzionali alcune operazioni fondamentali dell'analisi matematica, quali, per esempio, la derivazione, la convoluzione e le trasformate di Fourier e Laplace. Le distribuzioni possiedono numerose proprietà notevoli che ampliano le possibilità dell'analisi matematica classica: ad esempio ogni distribuzione è infinitamente differenziabile (in senso generalizzato), le serie convergenti di distribuzioni si possono differenziare termine a termine un numero arbitrario di volte, la trasformata di Fourier di una distribuzione temperata esiste sempre, e così via. L'uso delle distribuzioni estende in maniera essenziale l'ambito dei problemi considerati e, allo stesso tempo, automatizzando le operazioni elementari, conduce a semplificazioni notevoli. Scrivono Kolmogorov e Fomin [1974] introducendo le funzioni «generalizzate» o distribuzioni:

> In analisi si è portati spesso a considerare il termine «funzione» con un grado di generalità che varia secondo le questioni studiate [...]. A volte la nozione classica di funzione, anche se intesa nel senso più generale, cioè come una legge arbitraria che ad ogni valore di x appartenente

> al dominio di definizione della funzione fa corrispondere un numero $y = f(x)$, si rivela insufficiente [Kolmogorov e Fomin, 1974, p. 197].

Dopo aver illustrato con alcuni esempi le difficoltà che si presentano nella pratica matematica e fisica allorché ci si limita all'impiego delle funzioni ordinarie e motivando con ciò la necessità di una estensione del concetto di funzione, significativamente aggiungono:

> Sottolineiamo ancora una volta che l'introduzione delle distribuzioni è dovuta a problemi del tutto concreti e non alla semplice tendenza a estendere il più possibile le nozioni dell'analisi. In fisica, d'altronde, si è cominciato a far uso delle distribuzioni da molto tempo, in ogni caso da prima che fosse costruita la teoria matematica rigorosa delle distribuzioni [Kolmogorov e Fomin, 1974, p. 197].

È questo un fenomeno ricorrente nello sviluppo della matematica moderna, il cui rapporto con la fisica non si concretizza solo nel vasto campo di applicazioni fisiche dei risultati matematici; viceversa, numerose sono le teorie matematiche, talvolta anche quelle a prima vista più astratte, che trovano origini e motivazioni nella ricerca fisica.

2.2 HEAVISIDE E IL CALCOLO OPERAZIONALE

L'idea di «funzione impropria», concepita in modo euristico come un oggetto che può essere nullo tranne che in un punto e avere ugualmente integrale non nullo, entra stabilmente nel campo della matematica con le ricerche di O. Heaviside (1850-1925).

Heaviside, che di professione era un ingegnere telegrafico e telefonico, si occupò principalmente di elettromagnetismo. Nell'opera in tre volumi intitolata *Teoria elettromagnetica* [1893, 1899, 1912], egli riformulò la teoria elettrodinamica di Maxwell (scrisse, fra l'altro, le equazioni di Maxwell sotto forma vettoriale), e si dedicò allo studio sistematico della propagazione delle onde elettromagnetiche nel vuoto e via cavo.

Nell'articolo *Sugli operatori in fisica matematica* [1893], Heaviside, trattando un circuito elettrico, considera una carica data dall'espressione $Q(t) = Q_0 H(t)$, dove $H(t)$ è la «funzione di Heaviside»

$$H(t) = \begin{cases} 0 & \text{per } t < 0 \\ 1 & \text{per } t \geqslant 0 \end{cases} \tag{23}$$

e Q_0 una costante. Egli calcola l'intensità di corrente $I(t) = pQ(t)$, dove p è la notazione di Heaviside per l'operatore di derivazione d/dt, da lui utilizzato come un «simbolo» algebrico, con una tecnica euristica chiamata «calcolo operazionale». Per Heaviside la «funzione» $I(t)$ è tale che $\int_{-\infty}^{+\infty} I(t)\,dt = Q_0$.

> [...] È un nonsenso? È un risultato assurdo che indica la scarsa attendibilità della matematica operazionale o la necessità di qualche modifica della teoria? Niente affatto. Infatti Q è una funzione del tempo, e pQ ha il significato di tasso di variazione istantaneo di Q rispetto a t. Se, come nel caso in questione, Q è zero per $t < 0$ e costante per $t \geqslant 0$, pQ è zero eccetto che per $t = 0$, dove è infinito. Però l'integrale da $-\infty$ a $+\infty$ di pQ è Q_0. In altre parole, pQ è una *funzione completamente concentrata all'istante* $t = 0$ e tale che il suo integrale dia Q_0. È, per così dire, una funzione «impulsiva». L'idea di impulso è ben nota in meccanica ed è qui sostanzialmente la stessa. [...] La funzione

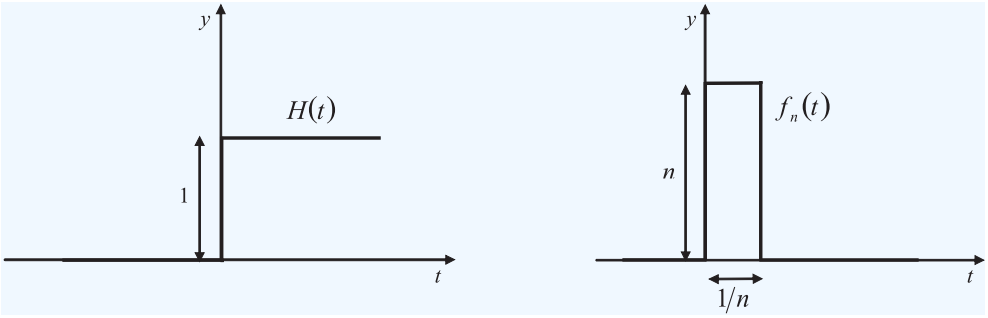


Figura 4.: Le «funzioni impulsive», introdotte in modo euristico come derivate simboliche di funzioni discontinue, sono considerate da Heaviside oggetti rappresentabili come limite di una successione di funzioni.

> pQ coinvolge gli ordinari concetti di derivata e integrale, *spinti al limite* (*pushed to their limit*) [corsivo nostro] [Heaviside, 1893, p. 14].

Le «funzioni impulsive», introdotte in modo euristico come derivate «simboliche» di funzioni discontinue, sono considerate da Heaviside, che qui riprende e sviluppa alcune idee di Fourier [1822] e di Kirchhoff [1882], oggetti ben determinati, che possono essere rappresentati analiticamente. Heaviside considera la successione di funzioni

$$f_n(t) = \begin{cases} n & \text{se } 0 \leqslant t \leqslant 1/n \\ 0 & \text{altrimenti} \end{cases} \tag{24}$$

la cui rappresentazione grafica è data nella figura 4. È «geometricamente evidente» che quando $n \to +\infty$ l'altezza della regione rettangolare rappresentata in figura cresce indefinitamente mentre la larghezza diminuisce in modo tale che l'area della regione sia sempre uguale a 1, cioè $\int_{-\infty}^{+\infty} f_n(t)\,dt = 1$. Si può quindi immaginare, continua Heaviside, una funzione limite cui tende $f_n(t)$ quando $n \to +\infty$. Questa funzione limite, detta «funzione impulsiva unitaria», coincide con la derivata «simbolica» $pH(t)$ della funzione di Heaviside. L'ingegnere inglese non entrò in maggiori dettagli in questi esempi. È comunque significativo osservare che le sue intuizioni possono essere rese rigorose: le espressioni

$$\frac{d}{dt} Q_0 H(t) = Q_0 \delta(t) \qquad \lim_{n \to +\infty} f_n(t) = \delta(t) \tag{25}$$

sono corrette nel senso distribuzionale.[1]

Negli anni Venti e Trenta diversi matematici, da N. Wiener (1926) e P. Lévy (1926) a B. Van der Pol (1929) e F. K. Niessen (1932), contribuirono a dare una sistemazione teorica al calcolo operazionale di Heaviside; nel 1937 G. Doetsch diede un assetto rigoroso alla teoria della trasformata di Laplace (per funzioni ordinarie), nel cui contesto si collocava anche il calcolo dell'ingegnere inglese. I tentativi di rendere rigoroso il calcolo operazionale si concentrarono però principalmente nella spiegazione della sua manipolazione algebrica di operatori differenziali, mentre i problemi connessi con le «funzioni impulsive» vennero per lo più trascurati. Poiché nella maggior parte delle applicazioni le funzioni impulsive entravano come $pH(t)$ e

1 Le nozioni e i procedimenti euristici usati da Heaviside sollevarono un grande scandalo fra i matematici, ed egli fu addirittura espulso dalla *Royal Society* di Londra per indegnità teorica.

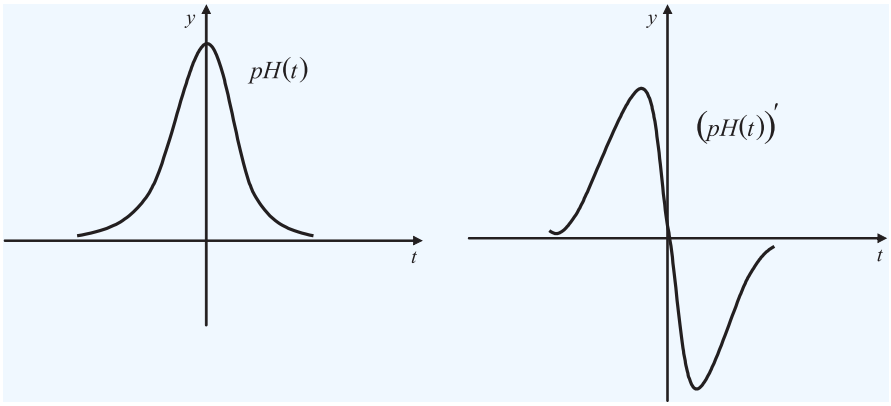


Figura 5.: Alcuni dei matematici impegnati nell'opera di rendere rigoroso il calcolo di Heaviside bandirono le «funzioni impulsive» trovandole «illegittime», mentre altri, fra cui Van der Pol, non si preoccuparono della debolezza dei loro fondamenti e le usarono senza particolari commenti.

scomparivano nel risultato, non veniva generalmente loro attribuito alcun particolare significato. In effetti le differenti sistemazioni del calcolo operazionale potevano spiegare il successo dei procedimenti di Heaviside solo nei casi in cui le funzioni impulsive non intervenivano. Alcuni dei matematici impegnati nell'opera di rendere rigoroso il calcolo di Heaviside bandirono tali funzioni trovandole «illegittime» [Doetsch, 1937, p. 57], mentre altri (fra i quali Van der Pol e Niessen) non si preoccuparono della debolezza dei loro fondamenti e le usarono senza curarsi del rigore. Van der Pol [1929] descrive la «funzione impulsiva unitaria» come segue:

$$\mathrm{pH}(t) = \begin{cases} 0 & \text{se } t \neq 0 \\ +\infty & \text{se } t = 0 \end{cases} \qquad \int_{-\infty}^{+\infty} \mathrm{pH}(t)\,dt = 1. \tag{26}$$

Afferma che tale oggetto può essere definito in modo equivalente come limite di una successione di gaussiane:

$$\mathrm{pH}(t) = \frac{n}{\sqrt{\pi}} e^{-n^2 t^2} \tag{27}$$

dove n è «un numero naturale molto grande». Egli inoltre dà la formula

$$\int_{-\infty}^{+\infty} f(t)\big(\mathrm{pH}(t)\big)'\,dt = -f'(0) \tag{28}$$

che dimostra attraverso un'integrazione per parti. Van der Pol illustra i suoi argomenti con la figura 5.

Per calcolare la trasformata di Laplace $\mathcal{L}[\mathrm{pH}(t)] = \int_0^{+\infty} \mathrm{pH}(t) e^{-st}\,dt$ della funzione impulsiva unitaria, Van der Pol procede come segue. Innanzitutto determina $\mathcal{L}[f_n(t)] = \int_0^{+\infty} f_n(t) e^{-st}\,dt$, dove $f_n(t)$ è definita dalla relazione (24); ottenuto $\mathcal{L}[f_n(t)] = \frac{1-e^{-s/n}}{s/n}$ mostra che $\lim_{n\to+\infty} \mathcal{L}[f_n(t)] = 1$ e conclude:

$$\mathcal{L}[\mathrm{pH}(t)] = \mathcal{L}\left[\lim_{n\to+\infty} f_n(t)\right] = \lim_{n\to+\infty} \mathcal{L}[f_n(t)] = 1 \tag{29}$$

La debolezza dell'impostazione di Van der Pol e Heaviside è dovuta essenzialmente ad una confusione circa le operazioni con i limiti doppi. Va però riconosciuto che la spregiudicatezza del loro modo di procedere coglie, in un certo senso, nel segno: bisogna sottolineare che tutti i risultati precedenti possono essere resi rigorosi nella teoria delle distribuzioni.

2.3 DELTA DI DIRAC

Il significato fisico delle funzioni «impulsive» o «improprie» è quello di fornire una rappresentazione di cariche e masse puntiformi. È degno di nota come fisici e matematici siano stati in grado di aggirare abilmente i problemi connessi con questi oggetti, in ambiti dove a noi, oggi, sembrano giocare un ruolo fondamentale.

Nel trattamento delle forze elettriche o gravitazionali, per esempio, la procedura tradizionale consisteva nel trattare per prime le cariche (o le masse) puntiformi e successivamente nel considerare distribuzioni continue di carica (o di massa), mediante l'introduzione di una opportuna funzione densità. In questo modo, però, era impossibile rappresentare in modo rigoroso linee e superficie cariche, come anche le cariche (le masse) puntiformi da cui si era inizialmente partiti. Questa contraddizione logica era il prezzo pagato dai fisici per la mancanza di una teoria delle distribuzioni adeguatamente sistemata.

La situazione che precedette la scoperta della teoria delle distribuzioni sembra per certi aspetti paradossale. Né i fisici né i matematici riconoscevano le funzioni improprie come parte della propria disciplina; i matematici, se le usavano, le consideravano una nozione fisica intuitiva priva di rigore, mentre i fisici le consideravano mere idealizzazioni matematiche prive di significato fisico. Maxwell, nel suo *Trattato sull'elettricità e il magnetismo* [1873], commenta la formula del potenziale di una carica puntiforme

$$V = \frac{e}{r} \qquad r = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} \tag{30}$$

con queste parole:

> L'espressione del potenziale V [data dalla formula (30)] vale nello spazio fuori da una superficie chiusa contenente il punto (a, b, c), ma non possiamo supporre, *eccetto che per scopi puramente matematici*, che questa sia la forma corretta del potenziale in corrispondenza del punto (a, b, c). Infatti, se così fosse, la forza elettrostatica in corrispondenza della carica dovrebbe essere infinita, per cui sarebbe necessaria una quantità infinita di lavoro per caricare un punto con una quantità finita di carica [corsivo nostro] [Maxwell, 1873, p. 81].

Per Maxwell le cariche puntiformi non erano altro che «piccoli corpi le cui dimensioni sono trascurabili rapportate alle distanze in gioco».

Insieme con l'elettrodinamica, la meccanica quantistica fu il campo di applicazione più importante delle funzioni impulsive prima della teoria delle distribuzioni. La funzione impulsiva unitaria è spesso chiamata «funzione δ» o «Delta di Dirac» dopo che il fisico Dirac la impiegò come potente strumento in meccanica quantistica.

P. A. M. Dirac (1902-1984), compiuti gli studi di *Electrical Engineering* all'Università di Bristol, seguì successivi studi di matematica applicata nella stessa università. Passato all'Università di Cambridge nel 1923, volse decisamente la sua attenzione alla fisica teorica. Nel 1930 fu pubblicata la prima edizione del trattato di Dirac sulla meccanica quantistica, *I princìpi della meccanica quantistica*, che sarebbe divenuto un classico. Dirac vi trasferì organicamente l'insieme delle sue ricerche e la sua visione generale della

meccanica quantistica. Modifiche rilevanti furono introdotte nella seconda edizione del 1935 e nella terza del 1947.[2]

L'interesse fondamentale, si potrebbe dire esclusivo, di Dirac è stato per la fisica teorica. La matematica, nell'opera di Dirac, non è mai un fine, bensì uno strumento che a volte egli stesso crea, incurante del rigore, sulla base di un intuito sicuro: è questo il caso della «funzione δ», che avrebbe trovato una sua rigorosa interpretazione nell'ambito della teoria delle distribuzioni.

Nella formulazione di Dirac della meccanica quantistica è spesso utile trovare una base ortonormale dello spazio delle funzioni d'onda in termini di autofunzioni di qualche operatore lineare, che rappresenta una determinata osservabile fisica; se lo spettro dell'operatore è discreto, la condizione di ortonormalità si scrive $\langle \xi | \xi' \rangle = \delta_{\xi\xi'}$, dove ξ e ξ' sono autofunzioni, $\langle \cdot | \cdot \rangle$ è la notazione di Dirac per il prodotto scalare e $\delta_{\xi\xi'}$ è il simbolo di Kronecker. Se però lo spettro dell'operatore è continuo, e quindi la base ha cardinalità più che numerabile, non è possibile esprimere la condizione di ortonormalità in termini di funzioni ordinarie.

Per avere un formalismo adeguato a superare questa difficoltà, Dirac introduce «una quantità $\delta(x)$ dipendente da un parametro x», che soddisfa le condizioni

$$\int_{-\infty}^{+\infty} \delta(x)\,dx = 1 \qquad \delta(x) = \begin{cases} 0 & \text{se } x \neq 0 \\ +\infty & \text{se } x = 0 \end{cases} \tag{31}$$

> Per avere un'idea dell'andamento della $\delta(x)$, consideriamo una funzione della variabile reale x che si annulli dovunque all'infuori che in un piccolo intervallo di lunghezza ε attorno all'origine $x = 0$, e che in esso sia tanto grande che il suo integrale esteso a tale intervallo risulti uguale a 1. L'esatta forma della funzione entro l'intervallo non ha importanza [...]. Allora il limite per $\varepsilon \to 0$ sarà $\delta(x)$ [Dirac, 1930, p. 80].

A Dirac era chiaro che la «funzione δ» non è una funzione di x secondo la definizione usuale: tale definizione, infatti,

> [...] richiede che una funzione abbia un valore definito in ciascun punto del suo dominio. La «funzione δ» è invece qualcosa di più generale, che noi possiamo chiamare «funzione impropria» per indicare la sua differenza da una funzione definita nella maniera usuale. Così $\delta(x)$ non è una quantità che possa essere usata in generale nell'analisi matematica come una funzione tradizionale, ma il suo uso deve essere ristretto ad alcuni semplici tipi di espressione per cui sia ovvio che non possa sorgere alcuna contraddizione [Dirac, 1930, p. 80].

La più importante proprietà della la «funzione δ» è espressa dalla seguente equazione:

$$\int_{-\infty}^{+\infty} \delta(x) f(x)\,dx = f(0) \tag{32}$$

in cui $f(x)$ è una generica funzione continua di x.

> Si può facilmente dimostrare la validità di questa equazione dalla precedente rappresentazione approssimata della $\delta(x)$. Il primo membro della (32) può dipendere solo dai valori di $f(x)$ molto vicini all'origine, cosicché possiamo sostituire $f(x)$ con il suo valore nell'origine, $f(0)$, senza apprezzabile errore. L'equazione (32) segue allora dalla prima delle due equazioni (31) [Dirac, 1930, p. 80].

2 Per le sue ricerche «nell'ambito della fisica atomica» Dirac condivise il premio Nobel per la fisica con E. Schrödinger nel 1933.

Effettuando una traslazione nella (32) deduce la formula

$$\int_{-\infty}^{+\infty} \delta(x-a)f(x)\,dx = f(a) \tag{33}$$

in cui a è un numero reale qualunque.

> Dunque il processo di moltiplicare una funzione di x per $\delta(x-a)$ e di integrarla poi su tutti gli x equivale a sostituire x con a. [...] Le equazioni (32) e (33) mostrano che, sebbene una funzione impropria non abbia essa stessa un valore ben definito, lo ha invece un integrale in cui essa si presenti nell'integrando come fattore. Nella teoria quantistica una funzione impropria apparirà soltanto in espressioni che alla fine saranno usate in un integrando. Perciò sarebbe possibile riformulare la teoria in una forma tale che le funzioni improprie apparissero solo in integrandi: si potrebbero allora eliminare completamente tali funzioni. L'uso di funzioni improprie non implica perciò alcuna mancanza di rigore nella teoria, ma è soltanto un formalismo conveniente che ci mette in grado di esprimere in forma concisa certe relazioni che noi potremmo riscrivere, se necessario, in una forma in cui non comparissero funzioni improprie; però ciò si potrebbe fare soltanto in maniera più involuta, che tenderebbe a oscurare il ragionamento [Dirac, 1930, p. 81].

Usando il formalismo della «funzione δ», Dirac può procedere con la formulazione della teoria quantistica. Nel caso in cui lo spettro dell'operatore lineare che rappresenta l'osservabile fisica in questione sia continuo, la condizione di ortonormalità delle autofunzioni si scrive

$$\langle \xi|\xi'\rangle = \delta(\xi-\xi') \tag{34}$$

È verosimile che Dirac abbia dato alla «funzione impropria» il nome «δ» proprio per questa analogia con la δ di Kronecker. Nella seconda edizione dei *Princìpi della meccanica quantistica*, Dirac propose una definizione equivalente di «funzione δ», che consiste nel rappresentarla come derivata $H'(x)$ della funzione di Heaviside $H(x)$. Dirac verifica che tale definizione è equivalente alla (33) sostituendo nel primo membro della (32) $H'(x)$ al posto di $\delta(x)$ e integrando per parti. Conclude che «la funzione δ appare ogni qualvolta si differenzi una funzione discontinua». Dirac conosceva sicuramente questa definizione e altre proprietà della «funzione δ» dai suoi studi universitari di *Electrical Engineering*.

Per la «funzione δ» Dirac è in grado di scrivere un certo numero di «equazioni elementari».

> Tali equazioni sono essenzialmente regole di calcolo per sviluppi algebrici in cui compaiono funzioni δ. Il significato di una qualunque di queste equazioni è che i suoi due membri danno risultati equivalenti qualora vengano usati come fattori in un integrando [Dirac, 1930, p. 82].

Esempi di tali equazioni sono:

$$\delta(x) = \delta(-x) \qquad f(x)\delta(x) = f(0)\delta(x) \qquad \delta(ax) = \frac{1}{|a|\,\delta(x)} \tag{35}$$

Per la derivata, che è «certamente una funzione ancora più discontinua e impropria della $\delta(x)$», Dirac, trovò

$$\delta'(x) = -\delta'(-x) \qquad x\delta'(x) = -\delta(x), \tag{36}$$

$$\int_{-\infty}^{+\infty} f(x)\delta^{(n)}(x-a)\,dx = (-a)^n f^n(a) \tag{37}$$

Nelle edizioni successive a questa lista vennero aggiunte altre formule. Così nella terza edizione (1947) si trova

$$\lim_{n\to+\infty} \frac{n}{\pi} \frac{1}{1+n^2x^2} = \delta(x), \tag{38}$$

$$\lim_{n\to+\infty} f_n(x) = \delta(x) \quad \text{dove} \quad f_n(x) = \begin{cases} n/2 & \text{se } |x| < 1/n \\ 0 & \text{altrimenti} \end{cases} \tag{39}$$

Tutto questo mostra l'abilità di Dirac nella manipolazione della «funzione δ»: tutti i risultati precedenti possono essere resi rigorosi nella teoria delle distribuzioni.

Naturalmente allo stesso Dirac era chiaro che la «funzione δ» non era una funzione in senso classico e, ciò che più importa, agiva come un operatore (più precisamente come un funzionale), che associa ad ogni funzione continua φ il numero $\varphi(0)$: la valutazione in 0. Occorsero però diversi e importanti contributi concettuali per formulare una definizione matematicamente corretta della «funzione δ» e delle sue derivate, che avrebbero trovato una rigorosa interpretazione nella teoria delle distribuzioni.

2.4 SOBOLEV E LA FONDAZIONE DELLE DISTRIBUZIONI

Gli strumenti matematici per realizzare una generalizzazione del concetto di funzione ordinaria in termini di funzionali erano stati scoperti fin dal 1910.

Il teorema di rappresentazione, di F. Riesz (1910), stabilisce infatti che se $\mathcal{H}$ è uno spazio di Hilbert, reale o complesso, e T un funzionale lineare e continuo su $\mathcal{H}$, allora esiste uno ed un solo $\alpha \in \mathcal{H}$ tale che

$$T(x) = \langle x, \alpha \rangle \quad \forall \alpha \in \mathcal{H} \tag{40}$$

In particolare, se $\mathcal{H} = L^2([-\pi, \pi], \mathbb{C})$, allora esiste una ed una sola funzione $g \in \mathcal{H}$ tale che

$$T(f) = \int_{-\pi}^{\pi} f(t)\, \overline{g(t)}\, dt \quad \forall f \in \mathcal{H} \tag{41}$$

La rappresentazione dei funzionali tramite un integrale del tipo (41) divenne, nelle mani di Sobolev [1936] e Schwartz [1950-1951], lo strumento centrale per la definizione delle distribuzioni *regolari* (cioè indotte da funzioni L^1_{loc}). Riesz, d'altra parte, mostrò che non tutti i funzionali potevano essere rappresentati in una forma del tipo (41) e J. Radon [1913] provò che tali funzionali potevano essere rappresentati usando una misura differente da quella di Lebesgue.

Fin dal 1910 erano dunque noti risultati che, implicitamente, offrivano la possibilità di una generalizzazione del concetto di funzione ordinaria mediante l'impiego di spazi di funzionali lineari e continui (duali di spazi di funzioni). Ciononostante i funzionali non vennero impiegati per questo scopo fino al 1936, quando Sobolev cominciò ad usare i funzionali come «funzioni generalizzate», al fine di estendere la classe delle soluzioni ammissibili del problema di Cauchy per l'equazione differenziale alle derivate parziali di tipo iperbolico. La ragione per cui la teoria delle funzioni generalizzate richiese un tempo relativamente lungo per essere sviluppata è da ricercarsi principalmente nell'assenza di forti motivazioni in questa direzione e nel fatto che le «funzioni improprie», la cui fondazione non rigorosa richiamava

l'attenzione sulla necessità di una teoria sistematica delle funzioni generalizzate, erano utilizzate soprattutto dagli ingegneri e, più tardi, dai fisici che non conoscevano adeguatamente l'analisi funzionale. Le motivazioni per una generalizzazione del concetto di funzione si affermarono decisamente negli anni Venti e Trenta del XX secolo in numerose branche della matematica e della fisica-matematica; fu solo quando problemi concreti di analisi fornirono motivazioni adeguate che i funzionali vennero riconosciuti dai matematici come la base per una tale generalizzazione. I problemi che diedero origine alla teoria delle distribuzioni non sorsero dunque dall'analisi funzionale, che pure rappresentò lo strumento per risolverli, ma emersero da questioni concrete di fisica-matematica.

I fondamenti della teoria delle distribuzioni furono posti dal matematico sovietico S. L. Sobolev (1908-1989) che, nel 1936, fu il primo a definire correttamente le distribuzioni come funzionali. Sobolev dedicò gran parte della sua vita scientifica allo studio sistematico delle equazioni differenziali alle derivate parziali ed alla risoluzione di problemi di fisica-matematica. I suoi speciali interessi sull'equazione delle onde e altre equazioni iperboliche riflettono il suo impiego all'Istituto Sismologico di Mosca durante gli anni 1929-1932. Nel 1933 Sobolev cominciò una serie di articoli sul problema di Cauchy per le equazioni iperboliche; questi studi lo condussero due anni più tardi alla generalizzazione del concetto di funzione ordinaria, che rende il suo lavoro così interessante dal nostro punto di vista. Il suo approccio al problema di Cauchy per le equazioni iperboliche continuò il lavoro di S. Bochner (1927) e J. Hadamard (1932), ma seguì linee nuove e più semplici. In una breve nota, intitolata *Il problema di Cauchy nello spazio dei funzionali* [1935], Sobolev accennò alla possibilità di ampliare la classe delle soluzioni ammissibili per il problema di Cauchy estendendo il problema a un opportuno spazio di funzionali. La nota, che non contiene dimostrazioni, si conclude con la prospettiva di un più dettagliato studio della materia. Queste ricerche vennero pubblicate l'anno successivo nell'articolo *Nuovo metodo di risolvere il problema di Cauchy per le equazioni lineari iperboliche in forma normale*, che ora esamineremo. Sobolev considera la seguente equazione iperbolica

$$Lu = \frac{\partial^2 u}{\partial t^2} - \sum_{j,k=1}^{N} a_{jk} \frac{\partial^2 u}{\partial x_j \partial x_k} + \sum_{j=1}^{N} b_j \frac{\partial u}{\partial x_j} + cu = f \tag{42}$$

dove i coefficienti a_{jk}, b_j e c sono numeri reali, la matrice (a_{jk}) è definita positiva ed il dato $(x,t) \mapsto f(x,t)$ e l'incognita $(x,t) \mapsto u(x,t)$ sono funzioni definite in $\mathbb{R}^N \times]0,+\infty[$.

Le equazioni iperboliche intervengono in molti modelli matematici di fenomeni di vibrazione e di propagazione ondosa. Il caso più semplice e importante è quello in cui (a_{jk}) è la matrice identità e gli altri coefficienti sono nulli: l'equazione ottenuta è detta *equazione delle onde*. Nel caso particolare di una sola variabile spaziale l'equazione delle onde è chiamata anche *equazione della corda vibrante* e la sua interpretazione è la seguente: il valore $u(x,t)$ è lo spostamento trasversale all'istante t del punto della corda che in condizioni di riposo occupa l'ascissa x.

Per semplicità, Sobolev considera l'equazione delle onde, detta anche *equazione di D'Alembert*,

$$Lu = \frac{\partial^2 u}{\partial t^2} - c^2 \Delta u = f \quad \text{in } \mathbb{R}^N \times]0,+\infty[\tag{43}$$

dove c è una costante positiva che, se x e t hanno i significati di variabile spaziale e di variabile temporale rispettivamente, ha la dimensione fisica della velocità. Per la (43) suppone $t > 0$ e considera il «problema di Cauchy in avanti», che consiste nell'imporre due condizioni iniziali:

$$u(\cdot, 0) = u_0 \qquad \frac{\partial u}{\partial t}(\cdot, 0) = u_1 \quad \text{in } \mathbb{R}^N \tag{44}$$

dove u_0 e u_1 sono due funzioni assegnate in $\mathbb{R}^N$. Sobolev comincia col presupporre $u_0 \in L^1_{\text{loc}}(\mathbb{R})$ e prova che la funzione

$$u(x,t) = \frac{1}{2}\big(u_0(x+ct) + u_0(x-ct)\big) \tag{45}$$

risolve l'equazione delle onde (43) in $\mathbb{R} \times \mathbb{R}$ con $f = 0$ e $n = 1$. Dimostra poi che, più in generale, della stessa proprietà è posseduta dalla funzione

$$u(x,t) = \varphi(x+ct) + \psi(x-ct) \quad x, t \in \mathbb{R} \tag{46}$$

dove $\varphi, \psi \in L^1_{\text{loc}}(\mathbb{R})$ sono funzioni arbitrarie. Come esempio prende $u_0 = H$ e ottiene

$$\begin{aligned} u(x,t) &= \frac{1}{2}\big(H(x+ct) + H(x-ct)\big) \\ &= \begin{cases} 0 & \text{se } x+ct < 0 \\ 1/2 & \text{se } x+ct > 0 \text{ e } x-ct > 0 \\ 1 & \text{se } x+ct < 0 \end{cases} \end{aligned} \tag{47}$$

> [...] Dunque u è costante in ciascuna delle tre regioni angolari di $\mathbb{R} \times]0, +\infty[$ individuate dalle semirette $x \pm ct = 0$ e quindi *già le derivate prime di* u *non sono più funzioni*. Ciò nonostante u risolve l'equazione: abbiamo dunque una soluzione discontinua. Il discorso va però approfondito, dato che occorre precisare in che senso tale u è la soluzione cercata; il problema più delicato consiste nell'attribuire un significato preciso alle condizioni di Cauchy quando la funzione che prendiamo in considerazione è irregolare [corsivo nostro] [Sobolev, 1936, p. 5].

Se u_0 è di classe C^1 non ci sono problemi: infatti anche la (45) è una funzione di classe C^1 e le (44) con $u_1 = 0$ sono soddisfatte in senso classico. Se invece u_0 è solo continua, la prima delle (44) è soddisfatta in senso classico mentre la seconda non ha per ora un significato preciso in quanto $\partial u/\partial t$ non è una funzione ordinaria. La stessa osservazione vale anche per la prima delle (44) se u_0 è solo localmente sommabile, almeno finché non si sceglie un'interpretazione precisa. Si presenta dunque l'opportunità di definire la nozione di *soluzione debole*, che in ipotesi opportune è equivalente a quella di soluzione classica.[3]

Per introdurre la nozione di soluzione debole, Sobolev esamina dapprima le funzioni da un altro punto di vista. Fissata un funzione reale u, studia come l'integrale del prodotto di u per la generica funzione reale v dipende da quest'ultima; studia cioè l'applicazione

$$v \mapsto \int_\Omega u(x)v(x)\,dx \tag{48}$$

3 Una funzione misurabile in un aperto $\Omega \subset \mathbb{R}^N$ è detta *localmente sommabile* in Ω quando essa è sommabile (cioè integrabile con integrale finito) su tutti i compatti inclusi in Ω. L'insieme delle funzioni localmente sommabili in Ω è denotato con $L^1_{\text{loc}}(\Omega)$.

dove Ω è un aperto di $\mathbb{R}^N$. Naturalmente occorre fare delle ipotesi su u e su v in modo che l'integrale del prodotto converga: solo dopo queste ipotesi ha senso il problema. Se $u \in L^\infty$, allora ha senso considerare tutte le v integrabili; se invece $u \in L^1$, è possibile lasciar variare v in L^∞; se $\Omega = \mathbb{R}^N$ e u è una funzione polinomiale, perché l'integrale converga è necessario imporre a v delle condizioni di annullamento all'infinito. In generale, continua Sobolev, se u è fissata in una certa classe abbastanza ampia, occorrerà supporre v variabile in un insieme relativamente ristretto in modo da compensare nel prodotto la grande generalità di u. Allora, se $u \in L^1_{loc}$, occorre restringere la classe delle funzioni v fino a supporre che esse siano a supporto compatto. Una condizione sufficiente per la convergenza dell'integrale del prodotto uv è che v sia limitata e a supporto compatto.[4]

Per procedere nella trattazione è dunque utile fissare per v una classe funzionale che possa compensare l'irregolarità di u: Sobolev suppone senz'altro che le funzioni v siano di classe C^∞ e a supporto compatto e denota con $C^\infty_0(\Omega)$ lo spazio vettoriale delle funzioni a valori reali $v \in C^\infty(\Omega)$, con Ω aperto di $\mathbb{R}^N$, il cui supporto è un sottoinsieme compatto di Ω. Stabilito ciò, l'applicazione da studiare è il funzionale

$$v \mapsto \int_\Omega u(x)v(x)\,dx \quad v \in C^\infty_0(\Omega) \tag{49}$$

ove u è fissata in $L^1_{loc}(\Omega)$. Segnaliamo che, quando le funzioni di $C^\infty_0(\Omega)$ sono viste come «punti» nel dominio del funzionale (49), esse vengono chiamate *funzioni test*.

La prima proprietà che Sobolev dimostra è la seguente: l'applicazione (49) è un *funzionale lineare*. Per introdurre la seconda dà la seguente definizione: siano (v_n) una successione di elementi dello spazio $C^\infty_0(\Omega)$ e $v \in C^\infty_0(\Omega)$; si dice che (v_n) *converge a v in* $C^\infty_0(\Omega)$, e si scrive $v_n \to v$, quando sono verificate le condizioni seguenti:

(*i*) esiste un compatto $K \subset \Omega$ che include i supporti di tutte le v_n;

(*ii*) per ogni operatore D di derivazione parziale di ordine $\geqslant 0$ qualunque si ha

$$\lim_{n\to+\infty} Dv_n(x) = Dv(x) \quad \text{uniformemente in } \Omega \tag{50}$$

Allora, riprendendo il discorso delle proprietà del funzionale (49) associato a una funzione u localmente sommabile, Sobolev prova che vale la proprietà seguente, che chiama *continuità*: se $v_n \to v$ in $C^\infty_0(\Omega)$, allora $\int_\Omega uv_n \to \int_\Omega uv$. Se (v_n) è nelle condizioni dette e se K è un compatto che include i supporti di tutte le v_n, si ha infatti

$$\left|\int_\Omega uv_n - \int_\Omega uv\right| \leqslant \int_K |u|\,|v_n - v| \leqslant \|u\|_{1,K}\,\|v_n - v\|_{\infty,K} \tag{51}$$

e l'ultimo membro è infinitesimo.

A questo punto Sobolev dà la seguente definizione: siano (u_n) una successione di funzionali lineari e continui in Ω e u un funzionale lineare e

4 Sia u una funzione, reale o complessa, misurabile in un aperto Ω contenuto in R^N. Il *supporto* di u è il sottoinsieme di Ω, denotato con $\operatorname{supp} u$ e caratterizzato dalla condizione seguente: un punto $x_0 \in \Omega$ non appartiene a $\operatorname{supp} u$ quando esiste un intorno I di x_0 incluso in Ω tale che $u(x) = 0$ quasi dappertutto in I.

continuo del tipo (49); si dice che (u_n) *converge a* u *in senso debole*, e si scrive ancora $u_n \to u$, quando

$$\lim_{n\to+\infty} \int_\Omega u_n v = \int_\Omega uv \quad \forall v \in C_0^\infty(\Omega) \tag{52}$$

L'idea di generalizzare il concetto di funzione ordinaria mediante l'impiego di funzionali lineari e continui su spazi di funzioni, e le definizioni di convergenza in $C_0^\infty(\Omega)$, di funzionale continuo e di convergenza debole sono identiche al metodo usato più tardi da Schwartz per definire le distribuzioni *regolari* (cioè indotte da funzioni L^1_{loc}) in $\mathcal{D}(\Omega)$, dove $\mathcal{D}(\Omega)$ $\left(= C_0^\infty(\Omega)\right)$ è la notazione di Schwartz per indicare le funzioni C^∞ a supporto compatto.

Dopo queste definizioni di analisi funzionale, Sobolev ritorna al problema originale (43) e (44). Per introdurre la nozione di soluzione debole, egli suppone dapprima che u e i dati siano funzioni abbastanza regolari in modo da poter prendere in considerazione il problema di Cauchy (44) per l'equazione (43) *in senso classico*, cioè punto per punto, e cerca di esprimere diversamente il fatto che u sia una soluzione. Per fissare le idee suppone u di classe C^2 nel chiuso $\mathbb{R}^N \times [0, +\infty[$ e i dati continui.

Sobolev suppone che u risolva la (43) e verifichi le condizioni (44), fissa ad arbitrio una funzione test $v \in C^\infty\left(\mathbb{R}^N \times [0, +\infty[\right)$, moltiplica l'equazione per v e integra per parti. In questo modo dimostra che u verifica la condizione

$$\begin{aligned}\iint_{\mathbb{R}^N\times]0,+\infty[} u\left(D_t^2 v - c^2\Delta v\right) dx\, dt = \iint_{\mathbb{R}^N\times]0,+\infty[} fv\, dx\, dt \\ + \int_{\mathbb{R}^N} \left(-u_0(x)D_t v(x,0) + u_1(x)v(x,0)\right) dx\end{aligned} \tag{53}$$

per tutte le $v \in C^\infty\left(\mathbb{R}^N \times [0, +\infty[\right)$ con supporto limitato.

Viceversa, ferme restando le ipotesi di regolarità fatte su u e i dati, suppone che la (53) valga per ogni v che verifica le condizioni dette. Prendendo in particolare $v \in C_0^\infty\left(\mathbb{R}^N \times [0, +\infty[\right)$, dunque con supporto incluso nel semispazio $t > 0$, vede che u risolve l'equazione delle onde nel senso classico, dato che u è regolare. In questo modo Sobolev prova che, se f è una funzione continua nel semispazio $t \geqslant 0$, u_0, u_1 due funzioni continue su $\mathbb{R}$, u una funzione di classe C^2 nel semispazio chiuso $t \geqslant 0$, allora u verifica l'equazione delle onde (43) e le condizioni iniziali (44) in senso classico se e solo se vale la formula (53) per tutte le funzioni test $v \in C_0^\infty\left(\mathbb{R}^N \times [0, +\infty[\right)$.

L'interesse della (53) è dovuto al fatto che, perché essa abbia significato, non sono necessarie le ipotesi della proposizione precedente, così che i dati e soprattutto la funzione u possono essere molto più generali. In altre parole, il vantaggio della nuova formulazione del problema è il seguente: per poter scrivere la (53) non sono necessarie ipotesi di regolarità su u, f, u_0 e u_1, dato che le derivate agiscono solo sulla funzione test v. Dunque la (53) può essere scritta in condizioni molto generali.

A questo punto Sobolev è in grado di dare la seguente definizione: siano f una funzione misurabile nel semispazio $t > 0$ e integrabile su tutti i suoi sottoinsiemi misurabili e limitati, e u_0 e u_1 due funzioni appartenenti a $L^1_{loc}(\mathbb{R}^N)$; una funzione u misurabile nel semispazio $t > 0$ e integrabile su tutti i suoi sottoinsiemi misurabili e limitati è detta *soluzione debole* del problema di Cauchy in avanti (44) per l'equazione (43) quando vale la (53) per tutte le funzioni test $v \in C_0^\infty\left(\mathbb{R}^N \times [0, +\infty[\right)$.

L'articolo di Sobolev si conclude con la dimostrazione del seguente teorema: se $u_0 \in L^1_{loc}(\mathbb{R})$, allora la funzione (45) è una soluzione debole del problema di Cauchy in avanti per l'equazione delle onde con $n = 1$ e con i dati f e u_1 nulli.

Se paragoniamo il lavoro di Sobolev con la successiva teoria delle distribuzioni di Schwartz, i punti in comune sono molteplici. Sobolev e Schwartz avevano essenzialmente lo stesso obiettivo: generalizzare il concetto di funzione ordinaria e alcune operazioni classiche del calcolo infinitesimale, estendendo le possibilità dell'analisi matematica tradizionale. Il metodo impiegato dai due matematici per realizzare tale generalizzazione è, nelle linee guida, fondamentalmente lo stesso, e si basa sull'impiego di funzionali lineari e continui su opportuni spazi di funzioni. Vedremo che il concetto di «funzionale» di Sobolev coincide con quello di «distribuzione regolare» di Schwartz. Anche la nozione di convergenza debole corrisponde nelle due esposizioni; perfino la nozione di convergenza introdotta da Sobolev nello spazio delle funzioni $C^\infty_0(\Omega)$ coincide con la nozione di Schwartz di convergenza in $\mathcal{D}(\Omega)$.

Così Sobolev è il fondatore della teoria delle distribuzioni, sebbene il nome «distribuzione» sia stato dato ai funzionali solo più tardi da Schwartz.

Ci sono però anche importanti differenze tra l'impostazione di Sobolev e quella di Schwartz. Una prima differenza sta nel fatto che nel lavoro di Schwartz ci sono concetti, nozioni e teoremi che non trovano riscontro nell'opera di Sobolev. È questo il caso del concetto, molto importante, di distribuzione *singolare*, mediante il quale il matematico francese è finalmente in grado di dare una sistemazione rigorosa alle «funzioni improprie» come la Delta di Dirac; anche le nozioni di distribuzione temperata, di distribuzione $\mathcal{L}$-trasformabile, di trasformata di Fourier e di Laplace in senso distribuzionale, di prodotto tensoriale e convoluzione per distribuzioni, nozioni introdotte da Schwartz nella *Théorie des distributions* [1950-1951], non hanno un parallelo nell'opera di Sobolev.

Un'altra differenza tra l'impostazione di Sobolev e di Schwartz è che Sobolev si servì delle distribuzioni essenzialmente come strumento per la soluzione di un problema specifico, e cioè quello di estendere la classe delle soluzioni ammissibili del problema di Cauchy per le equazioni differenziali alle derivate parziali. Schwartz sviluppò invece la teoria delle distribuzioni rendendola una teoria versatile, applicabile ed applicata alla definizione rigorosa e alla risoluzione di molti problemi di matematica e di fisica-matematica. Rispetto alle diverse esigenze che motivarono la nascita della teoria delle distribuzioni, Sobolev si colloca essenzialmente in una sola di queste linee-guida (le soluzioni generalizzate di equazioni differenziali), mentre Schwartz ha il merito di aver compiuto una sintesi di tutte le diverse motivazioni teoriche.

Così *Sobolev inventò le distribuzioni*, ma è *a Schwartz* che va riconosciuto il merito di aver creato la *teoria* delle distribuzioni.[5]

Sobolev proseguì le sue ricerche sulle equazioni differenziali. Nel corso dei suoi studi [1963; 1964], introdusse, fra l'altro, la nozione di «derivata de-

5 Nei testi russi e dell'Europa dell'Est sulla teoria delle distribuzioni il merito della creazione della teoria viene attribuito a Sobolev e il contributo di Schwartz viene per lo più sottaciuto; nei testi occidentali è il lavoro di Sobolev ad essere generalmente trascurato, e la scoperta delle distribuzioni viene attribuita esclusivamente a Schwartz.

bole» e di «spazio di Sobolev», che sono ancor oggi strumenti fondamentali per lo studio delle equazioni differenziali alle derivate parziali.

2.5 TEORIA DELLE DISTRIBUZIONI DI SCHWARTZ

La costruzione della teoria delle distribuzioni venne affrontata sistematicamente dal matematico francese L. Schwartz (n. 1915) fra il 1945 e il 1951.

Come Sobolev, anche Schwartz dedicò gran parte della sua vita scientifica allo studio delle equazioni differenziali alle derivate parziali ed alla risoluzione di problemi di analisi funzionale e di fisica-matematica. Le motivazioni che spinsero Schwartz alla creazione della teoria delle distribuzioni risiedono nella volontà di dare un assetto rigoroso alla categoria delle funzioni improprie, come la Delta di Dirac e le sue derivate, e al contempo di porre su nuove basi la teoria delle equazioni differenziali ordinarie e alle derivate parziali, completando, di fatto, il lavoro di Sobolev.

Nel 1945 Schwartz pubblicò il primo di una serie di quattro articoli, presentati fra il 1945 e il 1948, che contengono le sue idee generali sulle distribuzioni: distribuzioni regolari e singolari, distribuzioni temperate ed $\mathcal{L}$-trasformabili, trasformate di Fourier e di Laplace in senso distribuzionale, prodotto tensoriale, convoluzione ed equazioni differenziali per distribuzioni. Secondo quanto riferisce lo stesso Schwartz [1950-1951; 1974], egli, all'epoca della pubblicazione dell'articolo del 1945, non conosceva il lavoro di Sobolev; fu J. Leray, nel 1946, a richiamare la sua attenzione sul lavoro del matematico sovietico.

I risultati di Schwartz sulle distribuzioni vennero esposti in maniera organica nella monografia *Théorie des distributions*, in due volumi [1950-1951], che divenne immediatamente il trattato standard sull'argomento e che ora esamineremo nelle sue linee essenziali.[6]

Dopo aver menzionato l'«audace» calcolo simbolico di Heaviside e gli studi degli «ingegneri elettrici» ad esso ispirati, nella premessa al primo volume della *Théorie* Schwartz osserva che, dopo l'introduzione della «funzione» di Dirac,

> [...] le formule del calcolo simbolico sono diventate ancor più inaccettabili per il rigore dei matematici. Scrivere che la funzione di Heaviside, uguale a 0 per $x < 0$ e a 1 per $x \geqslant 0$, ha per derivata la funzione di Dirac $\delta(x)$, la cui stessa definizione è matematicamente contraddittoria, e parlare di derivate $\delta'(x)$, $\delta''(x)$, ... di questa funzione priva di esistenza reale, è oltrepassare i limiti che ci sono permessi [Schwartz, 1950-1951, p. 5].

Schwartz osserva inoltre che

> [...] per quanto riguarda le equazioni differenziali alle derivate parziali, l'opportunità di generalizzare le impostazioni tradizionali si rivela impellente nello studio di svariati fenomeni fisici, per i quali si impone la necessità di estendere la classe delle soluzioni ammissibili [Schwartz, 1950-1951, p. 6].

Schwartz denota con $\mathcal{D}(\Omega)$ lo spazio vettoriale, reale o complesso, delle funzioni $v \in C^\infty(\Omega)$, con Ω aperto di $\mathbb{R}^N$, il cui supporto è un sottoinsieme compatto di Ω. Introduce poi la seconda definizione. Siano (v_n) una successione di elementi dello spazio $\mathcal{D}(\Omega)$ e $v \in \mathcal{D}(\Omega)$.

6 Per la creazione della teoria delle distribuzioni Schwartz venne insignito nel 1950 della Medaglia Fields.

Si dice che (v_n) *converge a* v *in* $\mathcal{D}(\Omega)$, e si scrive $v_n \to v$, quando sono verificate le condizioni seguenti:

(*i*) esiste un compatto $K \subset \Omega$ che include i supporti di tutte le v_n;

(*ii*) per ogni operatore D di derivazione parziale di ordine $\geqslant 0$ qualunque si ha

$$\lim_{n\to+\infty} Dv_n(x) = Dv(x) \quad \text{uniformemente in } \Omega \tag{54}$$

A questo punto dà la seguente definizione: sia Ω aperto di $\mathbb{R}^N$; si chiama *distribuzione su* Ω ogni funzionale L lineare su $\mathcal{D}(\Omega)$ che sia anche continuo nel senso seguente:

$$\text{da } v_n \to v \text{ in } \mathcal{D}(\Omega) \quad \text{segue} \quad Lv_n \to Lv \tag{55}$$

Le definizioni precedenti di convergenza nello spazio delle funzioni test $\mathcal{D}(\Omega)\,\big(= C_0^\infty(\Omega)\big)$ e di continuità per un funzionale lineare su $\mathcal{D}(\Omega)$ sono identiche a quelle introdotte da Sobolev nel lavoro del 1936.

Lo *spazio delle distribuzioni su* Ω viene denotato da Schwartz con $\mathcal{D}'(\Omega)$. Inoltre, se $u \in \mathcal{D}'(\Omega)$ e $v \in \mathcal{D}(\Omega)$, ciascuno dei simboli

$$\langle u, v\rangle \qquad \int_\Omega uv \qquad \int_\Omega u \cdot v \tag{56}$$

denota il valore assunto in v dal funzionale u. Schwartz osserva subito il fatto seguente: due distribuzioni u e w sono uguali se e solo se

$$\int_\Omega uv = \int_\Omega wv \quad \forall v \in \mathcal{D}(\Omega) \tag{57}$$

Infatti i due membri denotano i valori che i funzionali u e w assumono nel punto v e due applicazioni sono uguali se e solo se esse assumono lo stesso valore in ogni punto del loro dominio comune.

Meno banale è invece osservare che il simbolo di integrale introdotto nella (56) *non è un integrale di Lebesgue* ed è perciò arbitrario e ingiustificato pretenderne le proprietà, anche se si può convenire di continuare a chiamarlo integrale. Ad esempio è completamente insensata la proprietà additiva. Quindi, continua Schwartz, sebbene più di una analogia conforti sulla scelta della notazione, il fatto segnalato può suggerire l'uso di simboli diversi dall'integrale, ad esempio il simbolo $\langle u, v\rangle$, nel quale la distribuzione u e la funzione test v sono separate da una virgola.

Schwartz usa la notazione appena introdotta nella definizione e nei risultati che seguono: siano (u_n) una successione di distribuzioni in Ω e u una distribuzione in Ω; si dice che (u_n) *converge a* u *nel senso delle distribuzioni*, e si scrive $u_n \to u$ in $\mathcal{D}'(\Omega)$, quando

$$\lim_{n\to+\infty} \int_\Omega u_n v = \int_\Omega uv \quad \forall v \in \mathcal{D}(\Omega) \tag{58}$$

e diciamo che la serie $\sum u_n$ converge a u nel senso delle distribuzioni quando converge a u nel senso delle distribuzioni la successione delle ridotte.

La nozione di convergenza nel senso delle distribuzioni coincide con la nozione di convergenza debole introdotta da Sobolev nello spazio dei funzionali. Schwartz dimostra poi il seguente «teorema di completezza»: sia (u_n) una successione di distribuzioni in Ω tale che il limite

$$\lim_{n\to+\infty} \int_\Omega u_n v \tag{59}$$

esista finito per ogni $v \in \mathcal{D}(\Omega)$; allora esiste una distribuzione u in Ω alla quale (u_n) converge nel senso delle distribuzioni.

A questo punto osserva che la nozione di convergenza nel senso delle distribuzioni e il «teorema di completezza» si estendono in modo naturale al caso in cui le successioni siano sostituite da oggetti più generali, ad esempio al caso in cui si abbiano famiglie dipendenti da un parametro $\varepsilon > 0$ e si faccia tendere ε a 0^+.

L'intenzione dichiarata da Schwartz nell'introdurre le distribuzioni è quella di generalizzare il concetto di funzione localmente sommabile. Schwartz dimostra subito che, fissata $u \in L^1_{loc}$, il funzionale

$$L(v) = \int_\Omega u(x)v(x)\,dx = \langle u, v\rangle \tag{60}$$

è una distribuzione. Ebbene, sarebbe comodo chiamarla ancora u e il condizionale è d'obbligo. Infatti, se w è un'altra funzione localmente sommabile, diversa da u, non sembra escluso che alle due funzioni u e w venga associata la stessa distribuzione. Se effettivamente ciò avvenisse, continua Schwartz, la convenzione di chiamare con lo stesso nome la funzione e la distribuzione associata creerebbe solo confusione. Fortunatamente le cose vanno nel modo migliore possibile. Schwartz prova infatti il risultato seguente: se le due distribuzioni associate a due funzioni dalla (60) sono uguali, allora sono uguali le due funzioni di partenza. In altre parole, vale il seguente «lemma fondamentale»: siano $u_1, u_2 \in L^1_{loc}$ tali che

$$\int_\Omega u_1 v = \int_\Omega u_2 v \tag{61}$$

per ogni $v \in \mathcal{D}(\Omega)$; allora $u_1 = u_2$ quasi dappertutto in Ω. Risulta allora ben giustificata la seguente convenzione: sia $u \in L^1_{loc}(\Omega)$; allora la distribuzione data dalla (60) è denotata ancora con u ed è chiamata distribuzione *regolare* su Ω indotta dalla funzione $u \in L^1_{loc}$. In base a questa convenzione ogni funzione localmente sommabile induce una distribuzione (regolare).

Ora, conoscere una distribuzione u, che è un funzionale, significa conoscere il valore che esso assume in corrispondenza ad ogni funzione test v, valore che Schwartz indica con $\langle u, v\rangle$ oppure, convenzionalmente, con $\int_\Omega uv$, anche se quest'ultimo simbolo non denota un integrale di Lebesgue. Dire che una distribuzione u è indotta da una funzione L^1_{loc} significa allora dire che il simbolo $\int_\Omega uv$ è effettivamente un integrale, precisamente l'integrale del prodotto uv, dove u è la funzione localmente sommabile che prende lo stesso nome della distribuzione u di partenza.

Ad esempio la formula

$$\int_{-\infty}^{+\infty} uv = \int_0^{+\infty} v \quad v \in \mathcal{D}(\Omega) \tag{62}$$

definisce una distribuzione u: si tratta della distribuzione (regolare) indotta dalla funzione di Heaviside. In generale il simbolo $\int_\Omega uv$ va interpretato alla lettera se la distribuzione è una funzione e resta comodo simbolo convenzionale in caso contrario. Ad esempio la formula

$$\int \delta v = v(0) \quad \forall v \in \mathcal{D}(\mathbb{R}^N) \tag{63}$$

definisce una distribuzione, detta Delta di Dirac, che però non è indotta da alcuna funzione localmente sommabile: dunque il simbolo di integrale è convenzionale.

Schwartz verifica che la Delta di Dirac non è indotta da alcuna funzione localmente sommabile nel modo seguente. Suppone per assurdo che δ appartenga a L^1_{loc}; usando la (63) e il «lemma fondamentale» con $\Omega = \mathbb{R}^N \setminus \{0\}$, ottiene $\delta(x) = 0$ quasi dappertutto in Ω. Deduce che, poiché l'insieme costituito dalla sola origine ha misura nulla, $\delta = 0$ in $\mathbb{R}^N$ così che l'integrale di δv si annulla per ogni $v \in \mathcal{D}$. Scelta allora $v \in \mathcal{D}$ tale che $v(0) = 1$, la (63) viene contraddetta.

Le distribuzioni che non sono indotte da funzioni L^1_{loc} sono chiamate da Schwartz distribuzioni *singolari*.

A questo punto Schwartz costruisce alcune successioni di funzioni che convergono nel senso delle distribuzioni. In particolare mostra che il fatto che la successione tenda a 0 puntualmente in $\mathbb{R}^N$ e diverga a $+\infty$ nell'origine non è necessario né sufficiente per la convergenza verso la Delta di Dirac, anche se tali condizioni sono soddisfatte in molti casi particolari. Per esempio, sia (r_n) una successione strettamente positiva e infinitesima e sia (u_n) una successione di funzioni non negative tali che per ogni n valgano le condizioni seguenti:

$$\lim_{n\to+\infty} \int u_n = 1 \tag{64}$$

$$u_n(x) = 0 \quad \text{quasi dappertutto per } |x| \geqslant r_n \tag{65}$$

Allora $u_n \to \delta$ in $\mathcal{D}'(\mathbb{R}^N)$. Basta infatti ragionare come sopra usando ora, come è lecito dato che $u_n \geqslant 0$, il «teorema della media pesata»: per ogni n esiste un punto $x_n \in B_{r_n}$, dove B_{r_n} è la palla di centro l'origine e raggio e raggio r_n, tale che

$$\int u_n v = \int_{B_{r_n}} u_n v = v(x_n) \int u_n \tag{66}$$

così che il limite vale $v(0)$, cioè proprio $\int \delta v$.

Un altro esempio è il seguente: sia $u \in L^1(\mathbb{R}^N)$ tale che $\int u = 1$; allora

$$\lim_{n\to+\infty} n^N u(nx) = \delta(x) \quad \text{in } \mathcal{D}'(\mathbb{R}^N) \tag{67}$$

Se $v \in \mathcal{D}(\mathbb{R}^N)$ si ha infatti

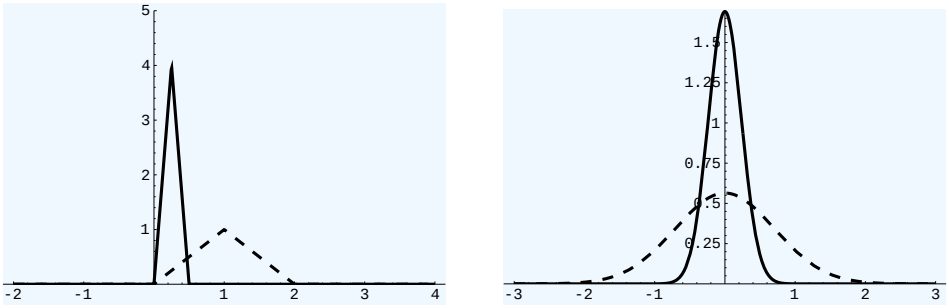
$$\begin{aligned} \lim_{n\to+\infty} \int n^N u(nx)v(x)\,dx &= \lim_{n\to+\infty} \int u(y)v(y/n)\,dy \\ &= \int u(y)v(0)\,dy = v(0) = \int \delta(y)v(y)\,dy \end{aligned} \tag{68}$$

grazie al «teorema di Lebesgue » («della convergenza dominata»), dato che

$$|u(y)v(y/n)| \leqslant \|v\|_\infty |u(y)|. \tag{69}$$

Ad esempio, nel caso $n = 1$, è nelle condizioni di ciascuno degli esempi precedenti la successione

$$u_n(x) = nu(nx) \quad \text{dove} \quad u(x) = \max(1 - |x - 1|, 0) \tag{70}$$



(a) (b)

Figura 6.: Le successioni $u_n(x) = nu(nx)$, con $u(x) = \max(1 - |x - 1|, 0)$ e rispettivamente $u(x) = \frac{1}{\sqrt{\pi}} e^{-x^2}$. Entrambe convergono alla Delta di Dirac nel senso delle distribuzioni sulla retta.

descritta nella figura 6. Schwartz nota esplicitamente che tutte le u_n sono nulle nell'origine e che la successione converge puntualmente alla funzione nulla. Ricordando che la convergenza quasi dappertutto e la convergenza in L^p sono compatibili (nel senso che se una successione converge contemporaneamente a w quasi dappertutto e a v in L^p, allora $w = v$), è chiaro che il test della convergenza quasi dappertutto, invece, non è utile per individuare il candidato quando si è interessati alla convergenza nel senso delle distribuzioni e può essere addirittura fuorviante, dato che una successione di funzioni può convergere quasi dappertutto a una funzione w e nel senso delle distribuzioni a una distribuzione diversa da w. Nel caso esaminato, w è la funzione nulla e la distribuzione limite è la Delta di Dirac, così che la convergenza quasi dappertutto non riveste alcun interesse.

A questo punto Schwartz affronta il problema di estendere, per quanto possibile, le operazioni elementari, note per le funzioni, al caso delle distribuzioni. Le prime operazioni sono quelle che conferiscono a $\mathcal{D}'(\Omega)$ una struttura di spazio vettoriale. Se u_1 e u_2 sono due distribuzioni e se c è uno scalare, la somma $u_1 + u_2$ e il prodotto cu_1 sono naturalmente i due funzionali

$$v \mapsto \int_\Omega u_1 v + \int_\Omega u_2 v, \quad v \mapsto c \int_\Omega u_1 v \tag{71}$$

dove v varia in $\mathcal{D}(\Omega)$, e si vede senza difficoltà che essi sono ancora distribuzioni. Con queste operazioni $\mathcal{D}'(\Omega)$ è uno spazio vettoriale, reale o complesso a seconda che si considerino distribuzioni reali o complesse, e la forma $\int_\Omega uv$ è bilineare. Dunque per il simbolo di integrale continua a valere una delle proprietà formali usuali.

Per quanto riguarda altre operazioni Schwartz si comporta in modo analogo. Tuttavia, dal momento che la loro introduzione è in generale meno indolore, scandisce con maggior cura il processo della definizione adottando la linea costituita dalle quattro tappe seguenti:

(*i*) vede come si comporta l'integrale rispetto all'operazione considerata nel caso delle funzioni ed evidenzia una proprietà degna di essere conservata al caso delle distribuzioni, proprietà che, in generale, sarà espressa da una formula il secondo membro della quale ha significato anche nel caso delle distribuzioni;

(*ii*) enuncia un lemma preliminare che afferma che il secondo membro della formula di cui sopra definisce, al variare della funzione test, una distribuzione;

(*iii*) dà la definizione in accordo con quanto stabilito dal lemma;

(*iv*) enuncia altre proprietà dell'operazione appena definita, con particolare riferimento a quelle che fanno intervenire la convergenza nel senso delle distribuzioni.

In generale si può dire che il problema della definizione può essere piuttosto delicato, in quanto occorre introdurre ipotesi opportune senza le quali il procedimento non avrebbe senso, ma che, in contrasto con questo fatto, le proprietà dell'operazione sono ottime e spesso di dimostrazione assai facile. Dunque il problema risiede spesso nella possibilità di estendere la definizione dell'operazione e non tanto nel fatto che le proprietà valide per le funzioni non si possano estendere alle distribuzioni.

Con alcuni esempi Schwartz [1950-1951, p. 34-36] mostra che non vi è alcuna definizione naturale di prodotto di due distribuzioni, né una definizione naturale delle composizioni $F \circ u$ e $u \circ F$ con u distribuzione, nemmeno quando F è una funzione polinomiale, per cui valgano buone proprietà rispetto alla convergenza nel senso delle distribuzioni.

Per quanto riguarda la definizione di prodotto di una funzione per una distribuzione per una funzione Schwartz si comporta come segue. Innanzitutto osserva che se $u \in L^1_{loc}(\Omega)$ e ψ è una funzione definita in Ω e se non sorgono problemi di integrabilità, il prodotto verifica l'uguaglianza

$$\int_\Omega wv = \int_\Omega u(\psi v) \quad \forall v \in \mathcal{D}(\Omega) \tag{72}$$

Per conservare questa proprietà anche nel caso in cui w sia il prodotto di ψ per una distribuzione u, occorre vedere innanzitutto se il secondo membro è ben definito per ogni $v \in \mathcal{D}$ anche quando u è una distribuzione e, per rispondere affermativamente, occorre come minimo sapere che $\psi v \in \mathcal{D}$: ebbene ciò e vero se ψ è una funzione di classe C^∞.

In altre parole, Schwartz mostra che vale il seguente lemma: siano $u \in \mathcal{D}'(\Omega)$ e ψ una funzione di classe $C^\infty(\Omega)$. Allora il funzionale w che a ogni $v \in \mathcal{D}(\Omega)$ associa il secondo membro della (72) è una distribuzione che rende vera la (72) stessa; tale distribuzione è denotata con il simbolo $u\psi$. Per esempio, se $\psi \in C^\infty(\mathbb{R})$ si ha $\psi\delta = \psi(0)\delta$. Infatti

$$\int (\psi\delta)v = \int \delta(\psi v) = (\psi v)\big|_{x=0} = \psi(0)v(0) = \psi(0)\int \delta v = \int \big(\psi(0)\delta\big)v \tag{73}$$

qualunque sia la funzione test $v \in \mathcal{D}(\mathbb{R}^N)$.

A questo punto Schwartz introduce il concetto di *supporto* di una distribuzione. Per le funzioni misurabili in un aperto Ω, e quindi in particolare per le funzioni $u \in L^1_{loc}(\Omega)$, la nozione di supporto è definita come segue: un punto $x \in \Omega$ non appartiene a $\operatorname{supp} u$ quando x ha un intorno in cui u è nulla quasi dappertutto. La definizione data da Schwartz nel caso delle distribuzioni estende quella nota nel senso che, quando viene applicata a una funzione localmente sommabile, fornisce il supporto già noto. Siano dunque Ω un aperto di $\mathbb{R}^N$ e $u \in \mathcal{D}'(\Omega)$; si dice che un punto non appartiene al supporto $\operatorname{supp} u$ di u quando esiste un aperto $\Omega' \subseteq \Omega$ intorno di x tale che la

restrizione $u|_{\Omega'}$ della distribuzione u all'aperto Ω' è nulla. Come esempio, Schwartz prova che

$$\operatorname{supp} \delta = \{0\} \tag{74}$$

Infatti la restrizione di δ al completamento dell'origine, che è aperto, è nulla, per cui nessun punto $x \neq 0$ appartiene al supporto; d'altra parte il punto 0 appartiene al supporto in quanto, preso ad arbitrio $\varepsilon > 0$, la distribuzione δ non è nulla in B_ε dato che esiste almeno una funzione $v \in \mathcal{D}(B_\varepsilon)$ tale che $v(0)$ è non nullo.[7]

Schwartz definisce poi le derivate direzionali di una distribuzione qualunque seguendo la stessa via usata nelle definizioni precedenti, scegliendo cioè una proprietà notevole soddisfatta nel caso delle funzioni. Se $u \in C^1(\Omega)$ e $v \in \mathcal{D}(\Omega)$, denotando con w la derivata $D_r u$ nella direzione r, dimostra che vale la formula

$$\int_\Omega w(x)v(x)\,dx = -\int_\Omega u(x)D_r v(x)\,dx \tag{75}$$

che è una conseguenza del «teorema della divergenza ».

Prova poi che vale il seguente lemma: siano Ω un aperto di $\mathbb{R}^N$, $u \in \mathcal{D}'(\Omega)$ e r un versore di $\mathbb{R}^N$; allora il funzionale w che a ogni $v \in \mathcal{D}(\Omega)$ associa il secondo membro della (75) è una distribuzione che rende vera la (75) stessa. Si può quindi dare le seguente definizione: nelle condizioni del lemma, la distribuzione w è denotata con $D_r u$ oppure con $\partial u/\partial r$ e viene detta *derivata (nel senso delle distribuzioni) di* u *rispetto alla direzione* r.

Come esempio, Schwartz prova che la derivata della funzione H di Heaviside è la distribuzione Delta di Dirac. Infatti, se $v \in \mathcal{D}(\mathbb{R})$, è possibile usare la formula di integrazione per parti

$$\int_{-\infty}^{+\infty} H'v = -\int_{-\infty}^{+\infty} Hv' = -\int_0^M v' = v(0) = \int_{-\infty}^{+\infty} \delta v \tag{76}$$

dove M è tale che $\operatorname{supp} v \subseteq [-M, M]$.

Non solo ogni distribuzione è derivabile in ogni direzione, addirittura infinite volte dato che il procedimento può essere iterato, ma le derivate hanno ottime proprietà, come la (75) e quelle che seguono, valide in totale assenza di ipotesi tecniche. Se Ω è un aperto di $\mathbb{R}^N$, $u \in \mathcal{D}'(\Omega)$ e $r \in \mathbb{R}^N$ è un versore, si ha

$$\frac{\partial u}{\partial r} = \sum_{j=1}^{N} r_j \frac{\partial u}{\partial x_j} \tag{77}$$

7 Schwartz introduce il concetto di restrizione di una distribuzione su Ω a un aperto $\Omega' \subseteq \Omega$ come segue. Innanzitutto considera il caso in cui u sia una funzione: se $v \in \mathcal{D}'(\Omega)$ e w è la restrizione di u a Ω', occorre presentare l'integrale su Ω' di wv come un integrale su Ω; definita la funzione $\tilde{v}$ (detta *prolungamento banale* di v) come

$$\tilde{v}(x) = \begin{cases} v(x) & \text{se } x \in \Omega' \\ 0 & \text{se } x \in \Omega \setminus \Omega' \end{cases}$$

si ha immediatamente che $\int_{\Omega'} wv = \int_\Omega u\tilde{v}$. Vale il seguente lemma: siano $u \in \mathcal{D}'(\Omega)$ e Ω' un aperto incluso in Ω; allora il funzionale w che a ogni $v \in \mathcal{D}'(\Omega)$ associa il secondo membro dell'equazione precedente è una distribuzione che rende vera l'equazione stessa. Nelle condizioni del lemma la distribuzione w è chiamata *restrizione di* u *a* Ω' ed è denotata con il simbolo $u|_{\Omega'}$

Inoltre, se la successione di distribuzioni (u_n) converge alla distribuzione u nel senso delle distribuzioni, allora per ogni operatore di derivazione D di qualunque ordine si ha

$$\lim_{n\to+\infty} Du_n = Du \quad \text{in } \mathcal{D}' \tag{78}$$

Nel secondo volume della *Théorie des distributions* [1950-1951] Schwartz, dopo aver dato le definizioni e i risultati fondamentali sulle trasformate di Fourier per le distribuzioni temperate, estende la teoria delle trasformate di Laplace al caso delle distribuzioni. Si dice che una distribuzione $u \in \mathcal{D}'(\mathbb{R})$ è $\mathcal{L}$*-trasformabile* quando $\operatorname{supp} u \subseteq [0,\infty[$ ed esiste un numero reale tale che la distribuzione

$$t \mapsto e^{-\lambda t}u(t) \tag{79}$$

appartenga a $\mathcal{S}'(\mathbb{R})$. Inoltre, se u è trasformabile, l'estremo inferiore $\lambda(u)$ dell'insieme dei valori $\lambda \in \mathbb{R}$ tali che valga la (79) è detto *ascissa di convergenza* dell'integrale di Laplace della distribuzione trasformabile u.[8]

Se u è una distribuzione trasformabile, allora per tutti i valori $s \in \mathbb{C}$ tali che $\operatorname{Re} s > \lambda u$ la *trasformata di Laplace* di u è definita da

$$\mathcal{L}[u] = \int_{-\infty}^{+\infty} u(t)e^{-st}\,dt, \tag{80}$$

il secondo membro della quale è a sua volta definito da

$$\int_{-\infty}^{+\infty} u(t)e^{-st}\,dt = \int_{-\infty}^{+\infty} u(t)e^{-st}\cdot e^{-(s-\lambda)t}\zeta(t)\,dt \tag{81}$$

dove $\lambda \in \mathbb{R}$ e $\zeta \in C^\infty(\mathbb{R})$ verificano le condizioni

$$\lambda(u) < \lambda < \operatorname{Re} s \quad \zeta(t) = 1 \quad \forall t \geqslant -1 \tag{82}$$

$$\zeta(t) = 1 \quad \forall t \leqslant -2 \tag{83}$$

Come esempio, Schwartz mostra che

$$\mathcal{L}[\delta] = 1 \quad \forall s \in \mathbb{C}. \tag{84}$$

8 Lo *spazio* $\mathcal{S}'(\mathbb{R}^N)$ *delle distribuzioni temperate* è introdotto da Schwartz come segue. Innanzitutto denota con $\mathcal{S}(\mathbb{R}^N)$ lo spazio delle funzioni $v \in C^\infty(\mathbb{R}^n)$ tali che

$$Dv(x) = o(|x|^{-m}) \quad \text{per } |x| \to +\infty$$

per ogni $m \in \mathbb{N}$ e per ogni derivazione D di ordine qualunque. Le funzioni di $\mathcal{S}(\mathbb{R}^N)$ sono dette *funzioni regolari a decrescenza rapida*. A questo punto Schwartz dà la seguente definizione: se P è una funzione polinomiale in n variabili, allora $P(D)$ è l'operatore differenziale lineare a coefficienti costanti ottenuto formalmente sostituendo nell'espressione $P(x)$ ogni x_j con l'operatore di derivazione parziale rispetto a x_j e l'eventuale termine costante c con l'operatore cI di moltiplicazione per c. Inoltre, si dice che una successione (v_n) di elementi di $\mathcal{S}(\mathbb{R}^N)$ *converge in* $\mathcal{S}(\mathbb{R}^N)$ a una funzione $v \in \mathcal{S}(\mathbb{R}^N)$ quando

$$P(x)Q(D)v_n(x) \to P(x)Q(D)v(x) \quad \text{uniformemente in } \mathbb{R}^N$$

per tutte le coppie di funzioni polinomiali P e Q. Lo spazio $\mathcal{S}'(\mathbb{R}^N)$ delle distribuzioni temperate è lo spazio associato alle funzioni test di $\mathcal{S}(\mathbb{R}^N)$. Si dà infatti la seguente definizione: una distribuzione $u \in \mathcal{D}'(\mathbb{R}^N)$ è detta *temperata* quando verifica la condizione

$$\lim_{n\to+\infty} \int_{\mathbb{R}^N} uv_n = 0$$

per ogni successione (v_n) di elementi di $\mathcal{D}(\mathbb{R}^N)$ che tende a 0 nel senso di $\mathcal{S}$. Lo spazio delle distribuzioni temperate su $\mathbb{R}^N$ viene denotato con $\mathcal{S}'(\mathbb{R}^N)$.

Infatti, siccome δ ha supporto compatto, è possibile scrivere direttamente

$$\mathcal{L}[\delta] = \int_{-\infty}^{+\infty} \delta(t) e^{-st}\, dt = e^{-st}\big|_{t=0} = 1 \tag{85}$$

Osserviamo esplicitamente che con quest'ultimo risultato, e con gli altri che lo precedono, il lungo percorso del pensiero matematico verso la rigorizzazione concettuale delle «funzioni improprie» può dirsi finalmente concluso. Tale sistemazione logica rappresenta il punto di arrivo di oltre cinquant'anni di ricerche, cominciate verso la fine dell'Ottocento da Heaviside con l'introduzione del «calcolo operazionale» e proseguite da Van der Pol con gli studi sulle «funzioni impulsive» e da Dirac con l'introduzione della «funzione δ».

Il secondo volume della *Théorie des distributions* si conclude con l'applicazione delle distribuzioni alle equazioni differenziali alle derivate parziali e con la discussione del caso ellittico, parabolico e iperbolico.

La monografia di Schwartz divenne immediatamente il lavoro standard sulla teoria delle distribuzioni. Dopo la pubblicazione della *Théorie des distributions*, Schwartz continuò a lavorare alla teoria delle distribuzioni e ad aggiornare il libro, che «resta ancora l'esposizione migliore e più completa della teoria [Treves, 1975, p. 465]». Schwartz applicò le distribuzioni alla teoria delle particelle elementari [1974] ed estese la teoria delle misure di Radon. Nelle mani di Schwartz e dei suoi allievi, la teoria delle distribuzioni si è rivelata uno strumento essenziale sia nell'analisi sia nelle sue applicazioni alla fisica. Secondo Dieudonné [1981, p. 231] «il ruolo di Schwartz nella teoria delle distribuzioni è molto simile a quello giocato da Newton e Leibniz nella storia del calcolo infinitesimale»; un ruolo di sistemazione, sia concettuale sia notazionale, di un algoritmo fra i più «potenti e versatili» dell'analisi contemporanea. (Per un resoconto più dettagliato del vasto lavoro matematico di Schwartz dopo il 1950 cfr. l'autobiografia di Schwartz [1974].)

CONCLUSIONI

L'evidenza più significativa che emerge da questo lavoro è l'importanza della presenza di motivazioni «concrete» per lo sviluppo delle idee matematiche: la matematica è sostanzialmente *un'attività di proposta e di soluzione di problemi*, teorici o pratici, puri o applicati.

Le ricerche dei matematici impegnati nel problema della fondazione del concetto di funzione e della nozione di distribuzione non sono avvenute sulla spinta di una sempre maggiore generalizzazione dei risultati o a causa di «astratte» questioni di rigore, ma sono state spesso motivate da problemi sorti dalla pratica matematica o da questioni di fisica-matematica. È un'evidenza che la fisica-matematica e più in generale la matematica applicata sono state una spinta eccezionale per lo sviluppo della matematica pura. Ne sono esempi emblematici il problema della corda vibrante, le ricerche di Fourier e le questioni che diedero origine alla teoria delle distribuzioni.

Uno degli esiti più interessanti della polemica sul problema della corda vibrante fu quello di concentrare l'attenzione dei matematici sul concetto di funzione. La fisica del problema della corda vibrante fu la molla che spinse Euler ad estendere la classe delle funzioni ammissibili in analisi e ad introdurre, accanto alle funzioni analitiche (che sono le «curve continue» secondo Euler), quelle continue regolari a tratti (le «discontinue» o «miste» nella terminologia euleriana).

Anche l'opera di Fourier, che impose una sostanziale revisione del concetto di funzione e comportò un ulteriore ampliamento della classe di funzioni ammissibili in analisi, si motiva inizialmente con questioni di natura fisica. Le stesse affermazioni sulla rappresentabilità in serie, che sembrano svolgersi su un terreno esclusivamente analitico, sono da Fourier motivate proprio dalla necessità di risolvere problemi fisici; senza queste, egli dice, non si potrebbe fare un passo in avanti nelle sue ricerche. La matematica deve trovare nella realtà esterna (fisica) stimoli e motivazioni: questa «concretezza» (e anche modernità, dopo più di centocinquant'anni) è l'indicazione che emerge con chiarezza dall'opera di Fourier.

Per quanto riguarda le origini della teoria delle distribuzioni, è significativo osservare che fin dal 1910 erano noti risultati che, implicitamente, offrivano la possibilità di una generalizzazione del concetto di funzione ordinaria mediante l'impiego di spazi di funzionali lineari e continui. Ciò nonostante, fu solo quando problemi concreti di analisi e di fisica-matematica fornirono, negli anni Venti e Trenta del XX secolo, motivazioni adeguate per una generalizzazione del concetto di funzione che i funzionali vennero riconosciuti dai matematici come la base per tale generalizzazione. I problemi che diedero origine alla teoria delle distribuzioni non sorsero dall'analisi funzionale, che pure rappresentò lo strumento per risolverli, ma emersero da questioni concrete di fisica-matematica. In fisica, d'altronde, si era cominciato a far uso delle distribuzioni prima che fosse costruita la teoria matematica rigorosa delle distribuzioni. È questo un fenomeno ricorrente nello sviluppo della matematica moderna, il cui rapporto con la fisica non si concretizza solo nel vasto campo di applicazioni fisiche dei risultati matematici; viceversa, numerose sono le teorie matematiche, talvolta anche quelle a prima vista più

astratte, che trovano origini e motivazioni nella ricerca fisica.

L'approfondimento storico che ha caratterizzato lo studio condotto per la stesura di questo lavoro si è rivelato di particolare interesse in quanto in grado di evidenziare la differenza che intercorre fra la matematica come sistema organicamente strutturato e come disciplina *in fieri*. I protagonisti della ricerca matematica sono uomini concretamente impegnati a «far matematica» in modi nuovi: tentano strade inesplorate, enunciano proposizioni, le saggiano, ritornano sugli enunciati, riprendono questioni apparentemente risolte per darne nuove soluzioni, con nuove prospettive. La matematica, nel corso del suo sviluppo storico, sembra molto più simile a una scienza sperimentale, che all'impalcatura ipotetico-deduttiva cui siamo abituati; quest'ultima appartiene alla fase di riorganizzazione logica delle teorie, più che al momento della scoperta vera e propria.

A

ANALISI NON-STANDARD

La prefazione alla prima edizione del libro *Non-standard Analysis* di Abraham Robinson inizia così:

> Nell'autunno del 1960 mi resi conto che i concetti e i metodi della Logica Matematica contemporanea potevano fornire un'intelaiatura adatta allo sviluppo del Calcolo differenziale ed integrale per mezzo di numeri infinitamente piccoli ed infinitamente grandi [...] La teoria risultante fu chiamata da me analisi non-standard perché coinvolge ed era, in parte, ispirata dai cosiddetti modelli non-standard dell'Aritmetica la cui esistenza era stata sottolineata per primo da T. Skolem [Robinson, 1966].

Ovviamente non è possibile descrivere qui i risultati di logica ai quali Robinson fa riferimento.[1]

Una domanda sorge spontanea: che ragione c'è di introdurre l'analisi non-standard? È forse «più potente» dell'analisi moderna? (Chiamiamo così, per intenderci, quella da Weierstrass [1872] in poi). È stato dimostrato che le due teorie sono equivalenti nel senso che ogni teorema di analisi non-standard ha un equivalente in analisi moderna e viceversa. La ragione è diversa e, forse, è racchiusa in alcune frasi dell'eminente logico Kurt Gödel, inserite nella prefazione alla seconda edizione del libro di Robinson[1974], che qui riportiamo parzialmente:

> [...] L'analisi non-standard spesso semplifica in modo sostanziale le dimostrazioni, non solo di teoremi, ma anche di risultati profondi. [...] Questi dati di fatto dovrebbero impedire una interpretazione sbagliata, piuttosto comune, dell'analisi non-standard, vale a dire che essa sia un tipo di stravaganza o di moda passeggera dei logici matematici. Niente potrebbe essere più lontano dalla verità. Piuttosto, ci sono buone ragioni per credere che l'analisi non-standard, in una versione o nell'altra, sarà l'analisi del futuro [Gödel, 1974].

Lo scopo di questa appendice è di illustrare brevemente le idee principali che hanno permesso di arricchire il campo dei reali con i numeri che Leibniz chiamava «ideali»: gli infinitesimi e gli infiniti. Nell'analisi moderna, infinitesimi ed infiniti sono funzioni che tendono a zero o all'infinito, rispettivamente. Per Leibniz e Robinson, come vedremo, essi sono invece numeri veri e propri; in essi non c'è nulla di «dinamico».

Ci si può chiedere a questo punto, quale sia la ragione per «reinventare» numeri infinitesimi ed infiniti dal momento che Leibniz e tutti gli studiosi del Calcolo differenziale usavano già tali entità ideali. La ragione è molto semplice: non si conosceva un modo rigoroso per introdurre quei numeri ed operare con essi; in ultima analisi, nonostante i molti successi della teoria leibniziana (si pensi alla meccanica dei corpi celesti), questo fatto significava una intrinseca limitazione alle possibilità di crescita ed espansione della teoria stessa.

1 Quest'appendice è basata sull'eccellente sintesi contenuta in [Pagani e Salsa, 1990, p. 518-524].

Può essere interessante vedere come nel 1629 Pierre de Fermat (le cui idee verranno riprese da Leibniz) utilizzi gli infinitesimi nella definizione di derivata, da lui introdotta come misura dell'inclinazione della tangente a una curva in un punto. Fermat considera una secante passante per due punti: il punto dato, e un altro punto distante da esso un infinitesimo h. E calcola la tangente (trigonometrica) della tangente (geometrica) come rapporto incrementale, in maniera simile a quella odierna. Per esempio, nel caso della funzione $f(x) = x^2$:

$$f'(x_0) = \frac{(x_0 + h)^2 - x_0^2}{h} = \frac{2x_0 h + h^2}{h} = 2x_0 + h = 2x_0 \tag{86}$$

Qui h viene considerato diverso da 0 quando lo si semplifica come divisore, ma uguale a 0 quando poi lo si elimina alla fine: un procedimento che non poteva non sollevare seri dubbi sulla sua consistenza. Sotto l'incalzare delle numerose critiche, gradualmente la teoria degli infinitesimi di Leibniz fu abbandonata per lasciare posto all'analisi moderna, per oltre un secolo.

Arriviamo dunque a Robinson. Un aspetto estremamente interessante emerso dalle scoperte di logica è che sono possibili moltissime estensioni non-standard dell'analisi moderna: Robinson elabora una di esse, che ha l'importante proprietà di essere collegata all'analisi moderna in modo tale da poter interpretare e trasferire ogni tipo di definizione o risultato dall'una all'altra. Non possiamo certo presentare la costruzione originale di Robinson; ci limitiamo alla costruzione di un modello non-standard dei numeri reali, che indicheremo con ${}^*\mathbb{R}$, nel modo più elementare possibile. Accenneremo poi alle definizioni non-standard di continuità, di derivata ed integrale.

Per capire l'idea base consideriamo la costruzione di Cantor dei numeri reali [Pagani e Salsa, 1990, p. 219]. In questa costruzione si considerano successioni fondamentali di numeri razionali (r_n), $r_n \in \mathbb{Q}$, quelle cioè che soddisfano la condizione di Cauchy. Il punto chiave della costruzione di $\mathbb{R}$ sta nell'assegnare una relazione di equivalenza nell'insieme delle successioni: due di esse, (r_n) e (s_n), sono ritenute equivalenti se $r_n - s_n \to 0$ per $n \to +\infty$. Successioni equivalenti individuano *lo stesso numero reale*. Come conseguenza, le successioni seguenti,

$$(0), \quad (2^{-n}) \quad \left(\frac{1}{n}\right) \quad \left(\frac{1}{n^2}\right) \quad \dots \quad \left(\frac{1}{n^k}\right) \tag{87}$$

equivalenti nel senso precisato sopra, individuano lo stesso numero reale: lo zero. D'altra parte, queste successioni risultano molto diverse tra loro per quanto riguarda la velocità con la quale tendono a zero. La relazione di equivalenza introdotta da Cantor non segnala tale differenza.

L'idea di Robinson è di partire da successioni di reali e di introdurre nel loro insieme una relazione di equivalenza, in modo tale da tener conto di tale diversità. Si potrebbe pensare di considerare come «numeri» le singole successioni oppure di ritenerle equivalenti se differiscono solo per i valori assunti su un insieme finito di interi. Questo modo di agire è giustificato dal fatto che, dopotutto siamo interessati al comportamento delle successioni per $n \to +\infty$. Per le classi di equivalenza così ottenute si definiscono somma e prodotto nella solita maniera:

$$[r_n] + [s_n] := [r_n + s_n] \qquad [r_n] \cdot [s_n] := [r_n \cdot s_n] \tag{88}$$

Ci si accorge però subito che non possono valere le ordinarie regole di calcolo. Infatti, ad esempio, $[0]$ è l'elemento neutro della somma e le classi individuate dalle successioni

$$r_n = 0, 1, 0, 3, 0, 5, \dots \qquad s_n = 0, 0, 2, 0, 4, 0, \dots \tag{89}$$

sono diverse da $[0]$. Tuttavia è evidente che $[r_n] \cdot [s_n] = 0$. Non vale quindi la legge di annullamento del prodotto. Occorre una relazione di equivalenza più «bilanciata» nel senso che non deve identificare «troppe» successioni (come quella di Cantor) altrimenti non si esce dal campo dei reali e al contempo non «troppo poche» altrimenti otterremmo un insieme sicuramente più grande di $\mathbb{R}$ ma nel quale non possiamo operare con le ordinarie regole di calcolo. Il primo passo e di operare una suddivisione. dei sottoinsiemi di $\mathbb{N}$ in «grandi» e «piccoli». A tale scopo introduciamo una funzione m definita sui sottoinsiemi di $\mathbb{N}$ con le seguenti proprietà:

(*i*) per ogni $A \subseteq \mathbb{N}$, $m(A) = 0$ oppure $m(A) = 1$;

(*ii*) $m(\emptyset) = 0$, $m(\mathbb{N}) = 1$;

(*iii*) se $A \subset \mathbb{N}$ è finito allora $m(A) = 0$;

(*iv*) se $A_1, A_2, \dots, A_k$ sono sottoinsiemi di $\mathbb{N}$ a due a due disgiunti allora

$$m\Big(\bigcup_{j=1}^{k} A_k\Big) = \sum_{j=1}^{k} m(A_j) \quad \text{(proprietà di additività)} \tag{90}$$

Una funzione come m si chiama *misura finitamente additiva* su $\mathcal{P}(\mathbb{N})$. L'esistenza di m (anzi di infinite m) è una conseguenza dell'assioma di scelta.

DEFINIZIONE A.1: Sia $A \subseteq \mathbb{N}$. Diciamo che A è *piccolo* se $m(A) = 0$, che A è *grande* se $m(A) = 1$.

Sia ora $\mathbb{R}^{\mathbb{N}}$ l'insieme delle successioni a valori reali. In $\mathbb{R}^{\mathbb{N}}$ introduciamo la seguente relazione di equivalenza che indichiamo con il simbolo $\overset{*}{\sim}$:

$$\{r_n\} \overset{*}{\sim} \{s_n\} \text{ se e solo se } m\{i \in \mathbb{N} : r_i = s_i\} = 1 \tag{91}$$

ovvero se l'insieme degli indici su cui coincidono è un sottoinsieme *grande* di $\mathbb{N}$.

DEFINIZIONE A.2: Si pone ${}^*\mathbb{R} := \mathbb{R}^{\mathbb{N}}/\overset{*}{\sim}$.

L'insieme ${}^*\mathbb{R}$ si chiama a volte insieme degli *iperreali*. I suoi elementi sono classi di equivalenza (rispetto a $\overset{*}{\sim}$) di successioni a valori reali, che indicheremo con il simbolo $\langle\cdot\rangle$. Useremo lettere greche per indicare iperreali; così, $\alpha \in {}^*\mathbb{R}$ significa che $\alpha = \langle r_n \rangle$ per qualche successione $r_n \colon \mathbb{N} \to \mathbb{R}$. Introduciamo una relazione d'ordine in ${}^*\mathbb{R}$, che continueremo ad indicare con il simbolo $<$: se $\alpha = \langle r_n \rangle$, $\beta = \langle s_n \rangle$ allora $\alpha < \beta$ se e solo se $m\{i \in \mathbb{N} : r_i < s_i\} = 1$. Le operazioni di somma e prodotto (indicate ancora con $+$ e $\cdot$) si definiscono ponendo

$$\alpha + \beta := \langle r_n + s_n \rangle \text{ e } \alpha \cdot \beta := \langle r_n \cdot s_n \rangle. \tag{92}$$

PROPOSIZIONE A.1: *La struttura* ${}^*\mathbb{R}$ *è un campo ordinato.*

Come conseguenza, in ${}^*\mathbb{R}$ valgono tutte le usuali regole di calcolo. In ${}^*\mathbb{R}$, $\mathbb{R}$ è immerso in modo naturale, nel senso che ${}^*\mathbb{R}$ contiene un campo ordinato *completo*, che per il teorema di isomorfismo [Pagani e Salsa, 1990, p. 86] possiamo senz'altro identificare con $\mathbb{R}$. È questo l'insieme degli elementi di ${}^*\mathbb{R}$ della forma ${}^*x := \langle x \rangle$, classe di equivalenza individuata dalla successione costante $r_n = x$ per ogni $n \in \mathbb{N}$. Scriveremo senz'altro $\mathbb{R} \subset {}^*\mathbb{R}$ e x al posto di ${}^*x$.

La proprietà che più ci interessa è però che ${}^*\mathbb{R}$ è *un'estensione propria* di $\mathbb{R}$, cioè $\mathbb{R} \subset {}^*\mathbb{R}$ (un *allargamento*, nella terminologia di Robinson). Siano infatti $\alpha = \langle \frac{1}{n} \rangle$ ed $x \in \mathbb{R}_+$. Allora $\alpha > 0$ e l'insieme

$$\{ i \in \mathbb{N} : \tfrac{1}{i} \geqslant x \} \tag{93}$$

è vuoto oppure costituito da un insieme finito di indici e pertanto è *piccolo* nel senso della definizione A.1. Ne segue che $0 < \alpha < x$; poiché x è arbitrario in $\mathbb{R}_+$ si ha che $0 < \alpha < x$ per ogni $x \in \mathbb{R}_+$. Ciò implica $\alpha \notin \mathbb{R}$; numeri come α si chiamano infinitesimi. Più precisamente:

DEFINIZIONE A.3: Si dice che $\varepsilon \in {}^*\mathbb{R}$ è *infinitesimo* se $|\varepsilon| < x$ per ogni $x \in \mathbb{R}_+$. Si dice che $\omega \in {}^*\mathbb{R}$ è *infinito* se $|\omega| > x$ per ogni $x \in \mathbb{R}_+$.

Un esempio di infinito è $\langle n \rangle$. Notiamo che $\beta = \langle \frac{1}{n^2} \rangle$ è infinitesimo e risulta $\beta < \alpha$, dove α è l'infinitesimo $\langle \frac{1}{n} \rangle$, come facilmente si verifica. Così pure $\langle n^2 \rangle$ è un infinito maggiore di $\langle n \rangle$.

Per indicare che ε è infinitesimo scriveremo $\varepsilon \approx 0$; l'unico infinitesimo reale è lo zero. Due numeri α e β si dicono infinitamente vicini, e si scrive $\alpha \approx \beta$, se $\alpha - \beta \approx 0$. È poi facile verificare che:

(*i*) se $\varepsilon \approx 0$, $\eta \approx 0$, $x \in \mathbb{R}$, allora $\varepsilon + \eta \approx 0$, $\varepsilon \cdot \eta \approx 0$, $\varepsilon \cdot x \approx 0$;

(*ii*) se $\varepsilon \approx 0$, $\varepsilon \neq 0$, allora $\dfrac{1}{\varepsilon}$ è infinito;

(*iii*) se ω e ν sono infiniti positivi, $x \in \mathbb{R}$, $x \neq 0$, allora $\omega + \nu$ e $\omega \cdot x$ sono infiniti;

(*iv*) se ω è infinito, allora $\dfrac{1}{\omega} \approx 0$.

È chiaro che ${}^*\mathbb{R}$ *non* può essere completo, altrimenti sarebbe isomorfo ad $\mathbb{R}$; infatti l'insieme degli infinitesimi non può avere estremo superiore (e inferiore) in ${}^*\mathbb{R}$, come facilmente si verifica.[2]

Né ${}^*\mathbb{R}$ può soddisfare la proprietà di Archimede: infatti se $x \in \mathbb{R}_+$ ed $\varepsilon \approx 0$, $\varepsilon > 0$, non esiste $n \in \mathbb{N}$ tale che $n\varepsilon > x$.

I numeri di ${}^*\mathbb{R}$ si possono suddividere in finiti ed infiniti.

DEFINIZIONE A.4: Il numero $\alpha \in {}^*\mathbb{R}$ è *finito* se esiste $r \in \mathbb{R}_+$ tale che $|\alpha| < r$.

Ora, tutti i numeri finiti sono della forma $x + \varepsilon$ con $x \in \mathbb{R}$ ed $\varepsilon \approx 0$; più precisamente vale il seguente teorema, fondamentale per lo sviluppo del calcolo differenziale ed integrale.

2 Sia E l'insieme degli infinitesimi. Per assurdo supponiamo che esista $\alpha \in {}^*\mathbb{R}$ tale che $\alpha = \sup E$. Si hanno due possibilità: $\alpha \approx 0$ oppure $\alpha \not\approx 0$. Se $\alpha \approx 0$, per ogni altro infinitesimo positivo ε si ha $\alpha + \varepsilon \approx 0$ e $\alpha + \varepsilon > \alpha$: contraddizione. Se $\alpha \not\approx 0$ ed ε è un infinitesimo positivo si ha $\alpha - \varepsilon > \eta$ per ogni $\eta \in E$: ancora contraddizione.

TEOREMA A.1: *Sia $\alpha \in {}^*\mathbb{R}$, α finito. Allora esiste uno ed un solo numero reale x tale che $x - a \approx 0$.*

DEFINIZIONE A.5: Il numero reale x si chiama *parte standard* di α e si indica con $\mathrm{st}(a)$.

DEFINIZIONE A.6: Se $x \in \mathbb{R}$, l'insieme dei punti della forma $x + \varepsilon$ con $\varepsilon \approx 0$ si chiama con il termine leibniziano *monade* (o anche *alone*) di x e si indica con $\mu(x)$.

Una visione suggestiva della retta iperreale ${}^*\mathbb{R}$, potrebbe essere lquella rappresentata nella figura 7.

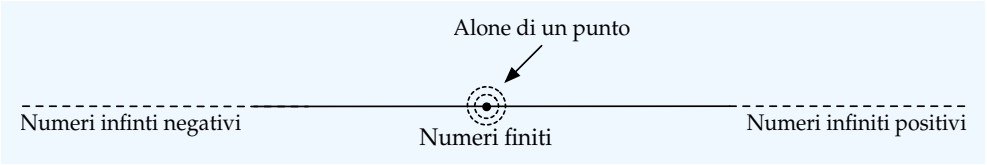


Figura 7.: La retta iperreale.

Ritorniamo ora al procedimento di estensione da $\mathbb{R}$ a ${}^*\mathbb{R}$. Esso si può applicare a qualunque sottoinsieme di $\mathbb{R}$, estendendolo a un sottoinsieme di ${}^*\mathbb{R}$: se $E \subset \mathbb{R}$, ${}^*E$ è definito dagli elementi $\alpha \in {}^*\mathbb{R}$ tali che, se $\alpha = \langle r_n \rangle$ allora

$$m\{i \in \mathbb{N} : r_i \in E\} = 1 \tag{94}$$

Così all'intervallo $[a, b]$ corrisponde l'intervallo ${}^*[a, b]$ che contiene la monade destra di a, quella sinistra di b e le monadi di tutti gli elementi di ${}^*(a, b)$.

Gli insiemi $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ si estendono a ${}^*\mathbb{N}$, ${}^*\mathbb{Z}$, ${}^*\mathbb{Q}$. In particolare ${}^*\mathbb{N}$ è costituito dagli interi di $\mathbb{N}$ e dagli interi infiniti: questi ultimi sono rappresentati da successioni del tipo $\langle n \rangle$, $\langle n^2 \rangle$, ...

I sottoinsiemi di ${}^*\mathbb{R}$ della forma ${}^*E$ con $E \subseteq \mathbb{R}$ si chiamano *insiemi standard*. Certamente essi non sono i soli sottoinsiemi di ${}^*\mathbb{R}$; ad esempio l'insieme degli infinitesimi non è standard; non possiamo però addentrarci più in profondità in questa breve introduzione.

Tramite la nozione di parte standard si possono introdurre le analoghe non-standard delle nozioni topologiche del capitolo terzo. Ad esempio, sia $A \subseteq {}^*\mathbb{R}$, A standard.

(*i*) A è aperto se per ogni $\alpha \in A$, α finito, si ha $\mu(\alpha) \subset A$;

(*ii*) A è chiuso se per ogni $\alpha \in A$, α finito, si ha $\mathrm{st}(\alpha) \subset A$;

(*iii*) A è compatto se ogni $\alpha \in A$ è finito e $\mathrm{st}(\alpha) \in A$.

Gli enti fondamentali dell'analisi sono le funzioni ed anch'esse si possono estendere con la procedura «$*$». Infatti, $f\colon E \to \mathbb{R}$, $E \subseteq \mathbb{R}$, si può estendere in modo naturale a ${}^*f\colon {}^*E \to {}^*\mathbb{R}$. Ad esempio $f(x) = x^2$ si estende a ${}^*f(\alpha) = \alpha^2$, $f(x) = \exp(x)$ si estende a ${}^*f(\alpha) = \exp(\alpha)$; come si vede la «corrispondenza» opera nello stesso modo anche sugli elementi non reali di ${}^*A$. Le funzioni ottenute in questo modo si chiamano *standard* e scriveremo senz'altro f invece di ${}^*f$.

Veniamo ora alle definizioni di continuità, derivata ed integrale per una funzione standard $f\colon {}^*(a, b) \to {}^*\mathbb{R}$.

DEFINIZIONE A.7 (continuità): Sia $x \in {}^*(a,b) \cap \mathbb{R}$; f si dice *continua* in x se

$$f(x+\varepsilon) \approx f(x) \quad \forall \varepsilon \approx 0 \tag{95}$$

ovvero se $\mathrm{st}[f(x+\varepsilon)] = f(x)$ per ogni $\varepsilon \approx 0$.

DEFINIZIONE A.8 (continuità uniforme): f è *continua uniformemente* in ${}^*(a,b)$ se $x', x\text{»} \in {}^*(a,b)$ e $x' \approx x\text{»}$ implica $f(x') \approx f(x\text{»})$.

DEFINIZIONE A.9 (derivata): Sia $x \in {}^*(a,b) \cap \mathbb{R}$; il numero reale $c \in R$ si dice *derivata* di f in x e si scrive $c = f'(x)$ se

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \approx c \quad \forall \varepsilon \approx 0 \tag{96}$$

ovvero se $\mathrm{st}\left[\dfrac{f(x+\varepsilon)-f(x)}{\varepsilon}\right] = c$ per ogni $\varepsilon \approx 0$.

DEFINIZIONE A.10 (integrale): Ci limitiamo per semplicità a funzioni continue in ${}^*[a,b]$. Sia $\omega \in {}^*\mathbb{N} \setminus \mathbb{N}$ un intero infinito e poniamo $\alpha_k = a + \frac{k}{\omega}(b-a)$, $k = 1, 2, \ldots, \omega$. I punti α_k costituiscono una suddivisione infinita di $[a,b]$, tale che $\alpha_k - \alpha_{k-1} = \frac{b-a}{\omega} \approx 0$. Poniamo

$$S_\omega(f) = \sum_{k=1}^{\omega} f(\alpha_k)(\alpha_k - \alpha_{k-1}) = \frac{b-a}{\omega} \sum_{k=1}^{\omega} f(\alpha_k) \tag{97}$$

Si osservi che $S_\omega(f)$ è somma di infiniti termini infinitesimi. Il *numero reale* I *si dice integrale di* f *in* $[a,b]$ *se*

$$S_\omega(f) \approx I \quad \forall \omega \in {}^*\mathbb{N} \setminus \mathbb{N} \tag{98}$$

ovvero se $\mathrm{st}[S_\omega(f)] = I$ *per ogni* $\omega \in {}^*\mathbb{N} \setminus \mathbb{N}$.

Come si vede l'operazione di parte standard sostituisce l'operazione di limite che qui non è necessaria.

Terminiamo questo flash sull'analisi non-standard con una dimostrazione allo scopo di dare un'idea, seppure minima, di ciò che Gödel intendeva con le parole riportate all'inizio. È la dimostrazione del teorema di Cantor-Heine: *ogni funzione continua su un compatto è uniformemente continua*.

Dimostrazione (non-standard). Sia $K \subset {}^*\mathbb{R}$ compatto e $f\colon K \to {}^*\mathbb{R}$ continua. Se x' e $x\text{»} \in K$ e $x' \approx x\text{»}$, sia $x = \mathrm{st}(x') = \mathrm{st}(x\text{»})$. Allora $x \in K$ essendo K compatto ed essendo f continua si ha

$$f(x') \approx f(x) \approx f(x\text{»}) \tag{99}$$

e il teorema è dimostrato. □

BIBLIOGRAFIA

Arbogast, Louis François Antoine

1791 «Mémoire sur la nature des fonctions arbitraires qui entrent dans les intégrales des équations aux différences partielles», *Bull. Sci. Math*.

Bell, Eric Temple

1937 *I grandi matematici*, Sansoni, Firenze.

Bernoulli, Daniel

1755 «Réflexions et éclaircissements sur les nouvelles vibrations des cordes», *Mém. Acad. Sci. Berl*.

Bottazzini, Umberto

1981 *Il calcolo sublime: storia dell'analisi matematica da Euler a Weierstrass*, Boringhieri, Torino.

1990 *Il flauto di Hilbert. Storia della matematica moderna e contemporanea*, UTET, Torino.

Bottazzini, Umberto, Paolo Freguglia e Laura Toti Rigatelli

1992 *Fonti per la storia della matematica*, Sansoni, Firenze.

Bourbaki, Nicolas

1939 *Les structures fondamentales de l'analyse. Théorie des ensembles (fascicule des résultats)*, Hermann, Parigi.

1960 *Elementi di storia della matematica*, Feltrinelli, Milano.

Boyer, Carl Benjamin

1968 *Storia della matematica*, ISEDI, Milano.

Carruccio, Ettore

1972 *Matematiche elementari da un punto di vista superiore*, Pitagora, Bologna.

Cauchy, Augustin-Louis

1821 *Cours d'analyse algébrique*, Hermann, Parigi.

1823 *Resumé des leçons sur le calcul infinitésimal*, Hermann, Parigi.

1844 «Mémoire sur les fonctions continues ou discontinues», *Bull. Sci. Math*.

Cavaillès, Jean

1938 *Remarques sur la formation de la théorie abstraite des ensembles*, Hermann, Parigi.

Condorcet, Jean-Antoine Caritat

1765 *Traité du calcul intégral*, Hermann, Parigi.

D'Alembert, Jean Baptiste

1749 «Recherches sur les vibrations des cordes», *Bull. Sci. Math*.

D'Alembert, Jean Baptiste

1752 «Addition au mémoire sur les vibrations des cordes», *Bull. Sci. Math*.

Dedekind, Richard

1872 *Continuità e numeri irrazionali*, a cura di O. Zariski, Stock, Roma.

1888 *Was sind und was sollen die Zahlen?*, Vieweg & Sohn, Berlino.

Demidov, Serge Sergeevich

1977 «Notion de solution des équations différentielles aux dérivées partielles», *Papers by Soviet Scientist at the* 15^{th} *Int. Congr. Hist. Sci*.

Dieudonné, Jean

1969 *Foundations of Modern Analysis*, Academic Press, New York.

1981 *History of functional analysis*, North Holland, Amsterdam.

Dirac, Paul

1930 *I Principi della meccanica quantistica*, Boringhieri, Torino.

Dirichlet, Peter Gustav Lejeune

1829 «Sur la convergence des séries trigonometriques», *Bull. Sci. Math.*

1837 «Über die Darstellung ganz willkürlicher Funktionen durch Sinus und Cosinusreihen», *Math. Ann*.

1889-1897 *Gesammelte Werke*, a cura di L. Kronecker, Keimer, Berlino.

Doetsch, Gustav

1937 *Theorie und Anwendung der Laplace-Transformation*, Springer, Berlino.

Dragoni, Giorgio, Silvio Bergia e Giovanni Gottardi

1999 *Dizionario biografico degli scienziati e dei tecnici*, Zanichelli, Bologna.

Euler, Leonard

1748 *Introductio in analysin infinitorum*, Bousquet, Losanna.

1750a *Institutiones calculi differentialis*, Bousquet, Losanna.

1750b *Sur la vibration des cordes*, Bousquet, Losanna.

Fourier, Jean Baptiste

1822 *Théorie analytique de la chaleur*, Hermann, Parigi.

Frajese, Attilio

1962 *Attraverso la storia della matematica*, UTET, Roma.

Gilardi, Gianni

1994 *Analisi tre*, McGraw-Hill, Milano.

Grattan-Guinness, Ivor

1972 *Joseph Fourier*, Cambridge Univeristy Press, Cambridge.

Hadamard, Jacques

1932 «Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques», *Comp. Rend. Acad. Sci.*

Hankel, Hermann

1870 «Untersuchungen über die unendlich oft oszillierenden und unstetigen Funktionen», *Math. Ann*, 20, p. 63-112.

Heaviside, Oliver
1893 «On operators in physical mathematics», *Proc. Roy. Soc.*
1912 «Electromagnetic Theory», *Proc. Roy. Soc.*

Hilbert, David
1899 *Fondamenti di Geometria*, Feltrinelli, Bologna.

Kirchhoff, Gustav Robert Georg
1882 «Zur Theorie der Lichtstrahlen», *Sitz. K. Preuss. Akad. Wiss.*

Kline, Morris
1972 *Storia del pensiero matematico*, Einaudi, Torino.

Kolmogorov, Andrej Nikolaevič e Sergei Fomin
1974 *Eléments de la théorie des fonctions et de l'analyse fonctionelle*, Mir, Mosca.

Kronecker, Leopold
1887 «Über den Zahlbegriff», *Math. Ann.*

Lacroix, Sylvestre-François
1810 *Traité du calcul différentiel et du calcul intégral*, Hermann, Parigi.

Lagrange, Joseph-Louis
1788 *Traité de Mécanique analytique*, Hermann, Parigi.
1797 *Théorie des fonctions analytique*, Hermann, Parigi.
1867-1892 *Œvres*, a cura di J. Serret, Hermann, Parigi.

Lanconelli, Ermanno
1994 *Lezioni di analisi matematica 1*, Pitagora, Bologna.
1995 *Lezioni di analisi matematica 2*, Pitagora, Bologna.
1996 *Primi elementi della Teoria degli Spazi di Hilbert*, Pitagora, Bologna.
1997 *Lezioni di analisi matematica 2 (Seconda Parte)*, Pitagora, Bologna.

Lebesgue, Henri
1906 *Leçons sur les séries trigonométriques*, Hermann, Parigi.

Lévy, Paul
1926 «Le calcul symbolique de Heaviside», *Bull. Sci. Math.*

Loria, Gino
1950 *Storia delle matematiche dall'alba della civiltà al tramonto del secolo XIX*, Hoepli, Milano.

Lützen, Jesper
1982 *The Prehistory of the Theory of Distributions*, Springer-Verlag, New York.

Maxwell, James Clerk
1873 *A Treatise on Electricity and Magnetism*, Springer-Verlag, Oxford.

Niessen, Karel Frederik e Balthazar van der Pol
1932 «Symbolic calculus», *Phil. Mag.*

Odifreddi, Piergiorgio

2000 *La matematica del Novecento*, Einaudi, Torino.

Pagani, Carlo Domenico e Sandro Salsa

1990 *Analisi Matematica 1*, Masson, Milano.

1991 *Analisi Matematica 2*, Masson, Milano.

Radon, Johann

1913 «Theorie und Anwendungen der absolut additiven Mengenfunktionen», *Sitz. Akad. Wiss. Wien. Math. Nat. Kl*.

Riesz, Frigyes

1910 «Untersuchungen über Systeme integrierbarer Funktionen», *Math. Ann*.

Robinson, Abraham

1966 *Non-Standard Analysis*, con prefazione di Kurt Gödel, North-Holland, Amsterdam.

Sachse, Arnold

1880 «Versuch einer Geschichte der Darstellung willkürlicher Funktionen einer Variablen durch trigonometrische Reihen», *Abh. Gesch. Math*.

Schwartz, Laurent

1945 «Généralisation de la notion de fonction, de dérivation, de transformation de Fourier, et applications mathématiques et physiques», *Ann. Univ. Grenoble Sect. Sci. Math. Phys*.

1947a «Théorie des distributions et transformation de Fourier», *Ann. Univ. Grenoble Sect. Sci. Math. Phys*.

1947b «Théorie générale des fonctions : théorie des distributions», *Ann. Univ. Grenoble. Sect. Sci. Math. Phys*.

1948 «Généralisation de la notion de fonction et dérivation : théorie des distributions», *Ann. Univ. Grenoble Sect. Sci. Math. Phys*.

1950-1951 *Théorie des distributions*, Hermann, Parigi.

1965 *Méthode mathématiques pour les sciences physiques*, Hermann, Parigi.

1974a «Applications des distributions à l'Étude des Particules Élémentaires en Mécanique Quantique Relativiste», *Acad. Sci. Paris*.

1974b «Notice sur les Travaux Scientifiques de Laurent Schwartz», *Acad. Sci. Paris*.

Sobolev, Sergej L'vovič

1935 «Le problème de Cauchy dans l'espace des fonctionelles», *Dokl. Acad. Sci. URSS*.

1936 «Méthode nuovelle à résoudre le problème de Cauchy pour les équations linéaires hyperboliques normales», *Mat. Sb*.

1963 «Applications of functional analysis in mathematical physics», *Amer. Math. Soc. Transl. Math. Mono*.

1964 *Partial differential equations of mathematical physics*, Oxford University Press, Oxford.

Treves, François

1975 *Basic Linear Partial Differential Equations*, Springer-Verlag, New York.

Van der Pol, Balthazar

1929 «On the operational solution of linear differential equations and an investigation of the properties of these solutions», *Phil. Mag*.

Vladimirov, Vasilij Sergeevič

1979 *Le distribuzioni nella fisica matematica*, Mir, Mosca.

Weierstrass, Karl

1872 «Über continuerliche Funktionen eines reellen Arguments», *Math. Ann*.

1886 «Abhandlungen aus der Funktionenlehre», *Math. Ann*.

Wiener, Norbert

1926 «The operational calculus», *Math. Ann*.

Zemanian, Armen

1965 *Distribution theory and transform analysis*, McGraw-Hill, New York.