

La delta di Dirac: storia di un mostro

Marco Bramanti
Dipartimento di Matematica - Politecnico di Milano

16 maggio 2021

Indice

1	Un primo sguardo sulla “funzione delta”	3
2	Chi ha inventato questo mostro?	
	Storia antica della delta	7
2.1	Esempi di “integrale contro la delta” nella matematica dell’800	8
2.2	Sorgenti concentrate e soluzioni fondamentali nella fisica matematica dell’800	11
2.3	Heaviside e le correnti impulsive	13
2.4	La delta “di Dirac”	15
3	La nascita della teoria delle distribuzioni di Laurent Schwartz	18
4	Le idee di base della teoria delle distribuzioni	22
4.1	Cos’è una funzione? La definizione puntuale	22
4.2	Le funzioni nella teoria moderna dell’integrazione	25
4.3	Le funzioni come funzionali	26
4.4	Nuovi funzionali. Le distribuzioni	30
4.5	Derivata di una distribuzione	32
4.6	Le distribuzioni come funzionali lineari continui*	34
5	Epilogo. Cosa abbiamo imparato?	37
5.1	Le trappole logiche (o psicologiche) dei vecchi racconti sulla delta	38
5.1.1	Impossibili scambi tra limiti e integrali	38
5.1.2	La necessità delle costruzioni matematiche	41
5.2	Cosa ci ha fatto guadagnare la teoria delle distribuzioni*	43
5.2.1	Equazioni alle derivate parziali	43
5.2.2	Successioni e serie di distribuzioni	44
5.2.3	Trasformata di Fourier	48
5.3	Matematica e conoscenza fisica	50

Introduzione

Questo scritto raccoglie il testo esteso preparato per una conferenza divulgativa che ho tenuto nel ciclo dei “Seminari di Cultura Matematica 2021”¹. L’argomento è un oggetto matematico ben preciso, molto specifico, la delta di Dirac, e il fatto di concentrarsi su un singolo oggetto così specifico è una scelta un po’ insolita come tema di una conferenza divulgativa. Ma ho proposto questo tema perché mi sembra che consenta di fare un racconto in qualche modo esemplare.

La “delta di Dirac”, infatti, è un oggetto matematico che è stato utilizzato sia in matematica che in fisica ormai da 200 anni, anche se il nome di Dirac è associato a questo oggetto a partire dal 1930. Prima della metà del ’90 nessuno sapeva dare una definizione di questo oggetto che non fosse palesemente autocontraddittoria. Eppure, lo si continuava a usare. Negli anni 1945-1950 il matematico Laurent Schwartz inventò la *teoria delle distribuzioni*, che tra le altre cose spiegava cosa fosse, come e perché si potesse usare quest’oggetto misterioso. Ma questa svolta non ha chiuso la partita, perché oggi la teoria delle distribuzioni fa parte del background di molti matematici e fisici, non di molti ingegneri. E ancora oggi in testi di fisica o ingegneria si parla spesso della delta di Dirac usando frasi, formule, “grafici esplicativi” che assomigliano a quelli che si trovavano nei lavori di 100 o 200 anni fa, e che non soddisfano i più elementari requisiti di coerenza logica che richiediamo alla matematica universitaria. E’ una situazione curiosa. In effetti la “delta” è un oggetto di cui non è facilissimo (ma, almeno oggi, neppure così difficile!) dare una descrizione rigorosa, eppure, anche se non la si descrive rigorosamente, non ci si vuole rinunciare, perché -tra le altre cose- rappresenta matematicamente delle situazioni fisiche idealizzate che sono estremamente utili: la densità di una massa unitaria concentrata in un punto dello spazio, un impulso elettrico che agisce per un tempo infinitamente breve spostando una carica elettrica misurabile, e varie altre situazioni del genere. La traduzione matematica “ingenua” di queste idee fisiche porta però a richiedere la validità di proprietà tra loro contraddittorie, “mostruose”. La storia della delta di Dirac è quindi la storia di un mostro di cui non ci si vuole liberare, perché si sa di averne bisogno. E’ la storia del rapporto intenso e tormentato tra intuizione e rigore matematico, tra le idee che pensiamo e le povere parole con cui a volte riusciamo a esprimere queste idee, tra gli utilizzatori della matematica e i matematici stessi. Questo racconto vuole costituire un’occasione di incontro tra le sensibilità diverse dei matematici teorici da una parte e dei fisici o degli ingegneri dall’altra.

Piano dell’articolo. Comincerò (§1) col presentare la delta rivolgendomi a chi non ne ha mai sentito parlare, e per far questo la presenterò nel suo significato fisico intuitivo, così come la si racconta di solito. Quindi (§2) racconterò le tappe principali della “storia antica” di quest’oggetto, riservando il §3 alla nascita

¹Seminario tenuto online il 5/5/2021 nel ciclo di Seminari di Cultura Matematica organizzati dal Laboratorio Efedisse del Dipartimento di Matematica del Politecnico di Milano. Alla pagina web dei seminari si può visionare il video della conferenza, v. http://fds.mate.polimi.it/?arg=divulgazione&id_pagina=226

della teoria delle distribuzioni. Nel §4 cercherò di spiegare, in modo divulgativo ma anche comunicando qualche idea matematica centrale, come la teoria delle distribuzioni abbia generalizzato il concetto di funzione e di derivata, e dato senso alla delta. Il §5 è un epilogo in cui cercherò di puntualizzare vari aspetti di quello che possiamo aver imparato da questa vicenda. L'articolo si chiude con qualche consiglio di lettura per chi vuole approfondire e i riferimenti bibliografici.

Avvertenza. Questo è uno scritto divulgativo che si rivolge a chi conosce almeno i primi elementi di calcolo differenziale e integrale per le funzioni di una variabile. Questo background basta per introdurre, in modo un po' semplificato ma abbastanza rigoroso, che cosa sia la delta di Dirac e quali siano le idee di base della teoria delle distribuzioni, che dà un senso alla delta nella matematica contemporanea. Spiegare le *motivazioni* per cui questo oggetto matematico è utile, come sia emerso storicamente dallo sviluppo della matematica e della fisica, così come fare qualche esempio dei benefici portati dalla teoria delle distribuzioni alla matematica contemporanea, coinvolge invece idee molto più ampie ed esigenti di matematica e di fisica. Per mantenere il discorso comprensibile a chi abbia solo i prerequisiti dichiarati, dovremo per forza accontentarci di qualche cenno a queste motivazioni, senza troppi dettagli. I due paragrafi contrassegnati con un * nel titolo (§ 4.6 e § 5.2) sono quelli più tecnici dal punto di vista matematico, in particolare nel § 5.2 si accenna anche alla trasformata di Fourier, che evidentemente esula dai “primi argomenti di calcolo differenziale e integrale”. L'omessa lettura di quei due paragrafi non pregiudica comunque la comprensione del contenuto principale di questo scritto. Così pure, in qualche nota a piè di pagina ho inserito qualche dettaglio più tecnico.

1 Un primo sguardo sulla “funzione delta”

Cominciamo a dare una prima idea del nostro oggetto misterioso, la “delta di Dirac”, per chi non ha mai sentito parlare, raccontandola “alla buona” come si fa di solito.

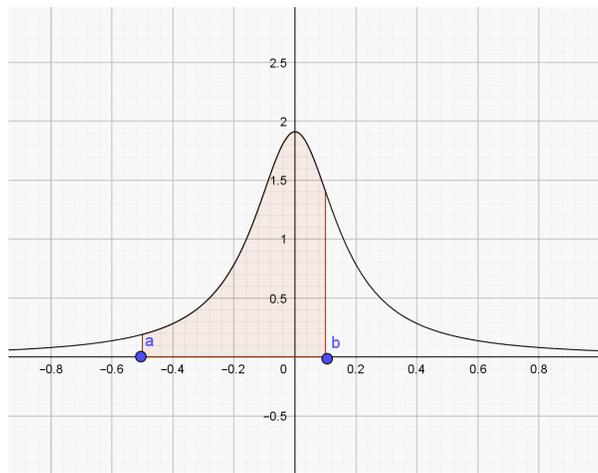
Consideriamo delle masse distribuite nello spazio; un discorso analogo si potrebbe fare con le cariche elettriche. Di solito in fisica si distinguono nettamente due situazioni diverse:

1. Oggetti puntiformi aventi ciascuno una massa ben definita (“punti materiali”); possiamo chiamare questa situazione *distribuzione discreta di masse*. In questo caso la distribuzione è specificata dicendo che in un certo punto P_1 si trova una massa m_1 , in un certo punto P_2 si trova una massa m_2 , e così via.

2. Oggetti materiali estesi, omogenei o non omogenei, come un filo, una sbarra, una superficie materiale, un solido, un fluido; possiamo chiamare questa situazione *distribuzione continua di masse*. Concentriamoci per semplicità sul caso unidimensionale: il nostro universo è una retta, e noi consideriamo ad esempio una o più sbarrette materiali, omogenee o non omogenee, distese lungo questa retta. La distribuzione di materia è allora descritta meglio introducendo una *funzione densità*. Se una sbarretta è omogenea, la sua densità (lineare,

(cioè unidimensionale) è semplicemente il rapporto $\rho = m/L$ tra la sua massa e la sua lunghezza. Nel caso più generale di una sbarretta non necessariamente omogenea, la densità sarà una funzione $\rho(x)$ del punto x della sbarretta, data dal rapporto tra la massa contenuta in un intervallino contenente x e la lunghezza Δx di quest'intervallino, al limite per $\Delta x \rightarrow 0$. La funzione $\rho(x)$ sarà ≥ 0 , identicamente nulla nelle regioni in cui non c'è materia. Se vogliamo sapere quanta massa si trova in una porzione $[a, b]$ della retta dovremo calcolare un integrale di ρ :

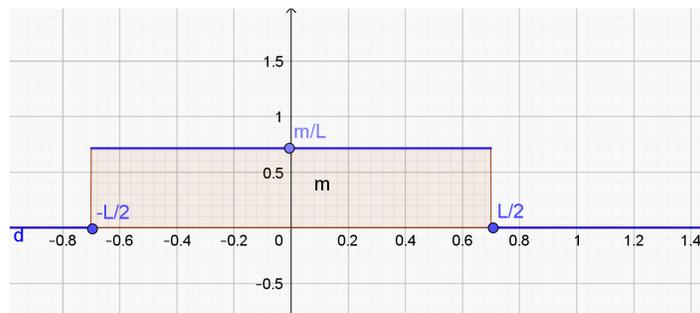
$$m_{[a,b]} = \int_a^b \rho(x) dx.$$



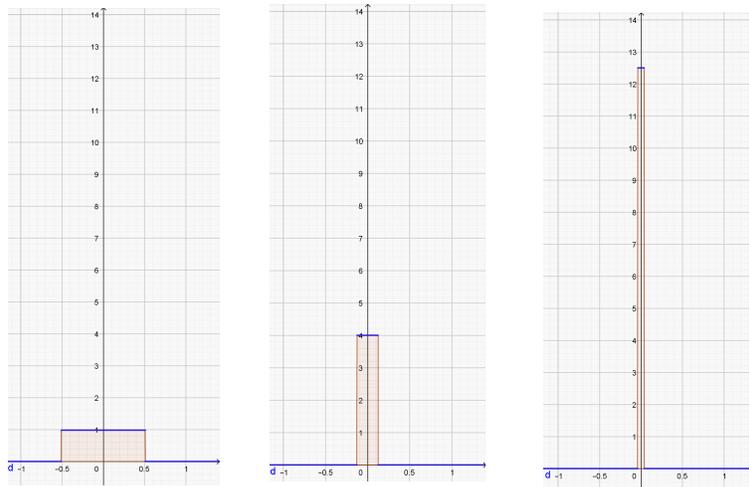
La massa totale distribuita sulla retta è

$$m = \int_{-\infty}^{+\infty} \rho(x) dx$$

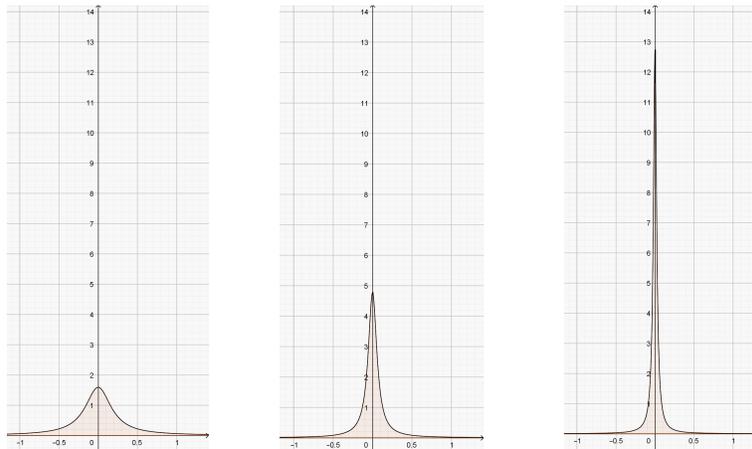
e si suppone che sia finita. Supponiamo per fissare le idee che la massa totale sia $m = 1$, quindi $\rho(x)$ è una funzione ≥ 0 , non necessariamente continua, con integrale 1. Anche la densità di una sbarretta omogenea (costante, uguale a $\rho = m/L$) può essere vista anche come una *funzione*, che vale m/L nei punti della sbarretta e zero altrove:



Ci chiediamo ora se sia possibile fare una sintesi dei due diversi tipi di distribuzioni di materia, quello discreto e quello continuo. Cos'è un punto materiale, dopo tutto? E' l'idealizzazione di una sbarretta talmente corta da avere lunghezza trascurabile, ma pur sempre massa positiva, ad esempio massa $m = 1$. Qual è la densità di questa distribuzione di materia? Se vogliamo rappresentare un punto materiale di massa $m = 1$, posto nell'origine, come caso limite di una distribuzione continua, possiamo pensare alla densità $\rho(x)$ come a una funzione nulla fuori da un piccolo intervallo di ampiezza L centrato nell'origine e avente un valore costante ρ in quell'intervallo, con $\rho \cdot L = 1$, quindi ρ molto grande, se L è molto piccolo.



Possiamo anche pensare a una curva $\rho(x)$ continua, con un grafico a campana (fisicamente: caso limite di una sbarretta non omogenea, con massa sempre più concentrata vicino all'origine), con un picco nell'origine:



Ognuna di queste funzioni $\rho(x)$ di cui facciamo il limite dovrebbe avere integrale 1 ed essere nulla fuori da un intervallo sempre più piccolo centrato nell'origine. Al limite, la funzione densità ρ , che da adesso in poi chiamiamo δ (cioè con questa lettera indichiamo l'ipotetica "densità limite" ottenuta con questo procedimento), dovrebbe soddisfare:

$$\begin{aligned}\delta(x) &= 0 \text{ per ogni } x \neq 0 \\ \delta(0) &= \infty \\ \int_{-\infty}^{+\infty} \delta(x) dx &= 1.\end{aligned}$$

Ora, il punto è che la prima e la terza di queste condizioni sono tra loro incompatibili: una funzione che vale zero tranne in un punto, ha certamente integrale zero, *qualunque valore assuma in quel punto*. Per capirlo, basta pensare all'integrale come *area sotto il grafico*: la prima condizione dice che l'integrale di δ è, al più, l'area di una semiretta, e questa vale zero. *Se vogliamo trattare sotto un unico quadro concettuale il caso di una distribuzione continua di massa, in cui il concetto di densità è adeguato, e il caso di una massa puntiforme, in cui il concetto di densità non è adeguato, siamo spinti a considerare delle "strane" funzioni densità con le proprietà di questa δ* . La "funzione delta" dovrebbe essere una funzione che ha le proprietà scritte qui sopra. Quindi è un oggetto impossibile, come un cavallo alato. Ma poiché questo "oggetto impossibile" risponde a esigenze molto sentite, in fisica, in ingegneria, anche in matematica teorica, non arrendiamoci alla prima difficoltà e andiamo avanti.

Riflettiamo ora su un'altra proprietà della delta che sarà centrale nel seguito. Torniamo a pensare a $\rho(x)$ come alla densità di una sbarretta (omogenea o non omogenea), con massa totale 1, che successivamente faremo di tendere ancora alla "densità del punto materiale", e chiediamoci: se $f(x)$ è una qualsiasi funzione continua, qual è il significato dell'integrale

$$\int_{-\infty}^{+\infty} f(x) \rho(x) dx ?$$

Per capirlo è più semplice ragionarci nel caso in cui ρ sia la densità di una sbarretta omogenea centrata in 0, quindi

$$\int_{-\infty}^{+\infty} f(x) \rho(x) dx = \frac{1}{L} \int_{-L/2}^{L/2} f(x) dx.$$

Il secondo membro dell'uguaglianza è la *media integrale* di f su un intervallino centrato in 0, e in particolare per $L \rightarrow 0$, cioè quando la densità è via via più concentrata in 0, se la funzione f è continua, questa media tende a $f(0)$:

$$\int_{-\infty}^{+\infty} f(x) \rho(x) dx = \frac{1}{L} \int_{-L/2}^{L/2} f(x) dx \rightarrow f(0) \text{ per } L \rightarrow 0.$$

Se invece di considerare la densità di una sbarra omogenea considerassimo la densità di una sbarra non omogenea di massa totale 1, via via più concentrata

vicino a 0, l'integrale $\int_{-\infty}^{+\infty} f(x) \rho(x) dx$ rappresenterebbe una sorta di “media pesata” dei valori di f , che tenderebbe ancora a $f(0)$ al tendere della densità $\rho(x)$ alla densità δ .

Otteniamo quindi, come caso limite, l'identità

$$\int_{-\infty}^{+\infty} f(x) \delta(x) dx = f(0), \quad (1)$$

per ogni funzione f continua. Quest'identità è altrettanto inverosimile quanto le proprietà della delta che abbiamo scritto sopra. Infatti il secondo membro dell'uguaglianza esprime una legge che agisce sulla funzione f in modo *locale*, cioè leggendo solo il valore di f in 0 (o per lo meno, dovendo essere f continua, leggendo i valori di f “in un intorno di 0 piccolo quanto si vuole”); invece il primo membro dell'uguaglianza esprime una legge che agisce sulla funzione f in modo *globale*, cioè leggendo i valori che f assume dappertutto (o per lo meno, in un intervallo di ampiezza positiva, se vogliamo che il valore dell'integrale non sia sempre nullo). E' proprio questa uguaglianza tra un oggetto globale e un oggetto locale il cuore della contraddizione.

L'identità precedente, traslata in un punto x_0 generico, diventa poi

$$\int_{-\infty}^{+\infty} f(x) \delta(x - x_0) dx = f(x_0). \quad (2)$$

Questa è una proprietà importante che si pretende valga per la delta. Spesso la si riscrive nella forma

$$\int_{-\infty}^{+\infty} f(x) \delta(x_0 - x) dx = f(x_0),$$

equivalente alla prima se si immagina δ come una funzione simmetrica pari. Come vedremo tra poco, quest'identità è esattamente la forma in cui storicamente la delta ha fatto il suo ingresso in società.

Si noti l'analogia tra questa formula e quella che vale in algebra lineare (e che, a differenza di quella qui sopra, non ha niente di misterioso): se $\{\delta_{ij}\}_{i,j=1}^n$ è la “delta di Kronecker” che vale 1 per $i = j$ e 0 per $i \neq j$, allora per ogni funzione $f(j)$ definita per $j = 1, 2, \dots, n$ si ha

$$\sum_{j=1}^n f(j) \delta_{ij} = f(i).$$

Così come δ_{ij} è diversa da zero solo per $i = j$, la $\delta(x - y)$ è diversa da zero solo per $x = y$.

2 Chi ha inventato questo mostro?

Storia antica della delta

Vediamo una carrellata di autori che a partire dall'800 hanno utilizzato in modo più o meno esplicito la funzione delta, in problemi fisici o matematici.

2.1 Esempi di “integrale contro la delta” nella matematica dell’800

Come anticipato, la funzione delta ha fatto la sua prima comparsa proprio attraverso identità integrali del tipo (2).

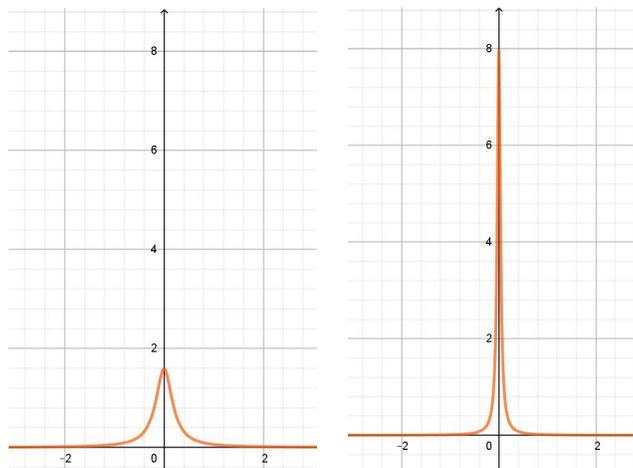
Parliamo di **Augustin-Louis Cauchy (1789 – 1857)**. Nel suo lavoro del **1815** (ma pubblicato nel 1827, [4]) *Théorie de la propagation des ondes à la surface d’un fluide pesant d’une profondeur indéfinie* (più noto col titolo breve *Memoria sulla teoria delle onde*), nella Nota XVIII, intitolata “Sur les intégrales définies singulières et les valeurs principales des integrales indéterminées”, scrive l’identità

$$F(x) = \frac{1}{\pi} \int_{x-\varepsilon}^{x+\varepsilon} F(y) \frac{\alpha}{\alpha^2 + (y-x)^2} dy$$

“dove ε e α sono numeri infinitamente piccoli”. La famiglia di funzioni

$$\delta_\alpha(y) = \frac{1}{\pi} \frac{\alpha}{\alpha^2 + y^2}, \quad (3)$$

oggi note in Analisi come “nucleo di Cauchy” (o, in Probabilità, “distribuzione di Cauchy”), hanno tutte integrale 1 e, per α sempre più piccolo, sono delle “campane” sempre più concentrate attorno a 0.



Grafici delle funzioni $\delta_\alpha(x)$ con $\alpha = 1/5$ e $\alpha = 1/25$

Quando il parametro α è molto piccolo, la curva a campana è molto concentrata, e quasi tutta la sua area è effettivamente compresa in un piccolo intorno di x . Quindi il fatto che gli estremi di integrazione siano $x - \varepsilon, x + \varepsilon$, “con ε infinitesimo”, non cambia molto le cose rispetto a considerare l’integrale su tutto \mathbb{R} . Il punto chiave è invece il ruolo del parametro α : Cauchy sta rappresentando un oggetto simile alla “funzione delta”, dicendo che è una funzione a campana dove il parametro α “è infinitamente piccolo”. Un modo ottocentesco per dire che quella funzione è il limite di una successione di curve a campana sempre

più concentrate. Un modo un po' insolito, non solo per noi, ma anche in un certo senso per Cauchy, che pochi anni dopo, nel 1821, col suo *Cours d'Analyse de l'École Royale Polytechnique*, avrebbe introdotto la nozione di *limite*, destinata a soppiantare il linguaggio degli infinitesimi. Questa (1815) è dunque, probabilmente, la prima comparsa della delta nella storia.

Passiamo ora a **Jean Baptiste Joseph Fourier (1768 – 1830)**. Nella sua opera *Théorie analytique de la chaleur*, del **1822**, in cui introduce gli sviluppi in serie trigonometriche per rappresentare le soluzioni dell'equazione di diffusione del calore (quelle che oggi chiamiamo “serie di Fourier”), a un certo punto l'autore introduce la *funzione impulsiva* δ (v. [9, pp.207-8], [18, p.112]). Partendo dall'identità che assegna una funzione

$$f : [-\pi, \pi] \rightarrow \mathbb{R}$$

come somma di una certa serie trigonometrica (la serie di Fourier di f , appunto) i cui coefficienti sono opportuni integrali, e rigirando quell'identità con passaggi illeciti², ottiene queste uguaglianze:

$$f(x) = \int_{-\pi}^{\pi} \delta(\alpha - x) f(\alpha) d\alpha \quad (4)$$

dove

$$\delta(x) = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right) \text{ per } x \in [-\pi, \pi]. \quad (5)$$

Fourier non usa il simbolo $\delta(x)$ ma dice a parole che l'espressione scritta nella (5) (calcolata in $(\alpha - x)$) è “una funzione di x e α tale che, se moltiplicata per

²Riporto questi passaggi per il lettore interessato. Partendo dallo sviluppo in serie di Fourier di una funzione f :

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx) \text{ dove:}$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ny) dy; \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ny) dy,$$

possiamo scrivere:

$$f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) dy + \sum_{n=1}^{\infty} \cos(nx) \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ny) dy + \sin(nx) \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ny) dy$$

(scambiando spensieratamente la serie con gli integrali)

$$= \int_{-\pi}^{\pi} f(y) \left\{ \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{\infty} \cos(nx) \cos(ny) + \sin(nx) \sin(ny) \right\} dy$$

$$= \int_{-\pi}^{\pi} f(y) \left\{ \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{\infty} \cos n(x - y) \right\} dy = \int_{-\pi}^{\pi} \delta(x - y) f(y) dy, \text{ con:}$$

$$\delta(x) = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right).$$

una qualsiasi funzione $f(x)$ e integrata tra $-\pi$ e π , dà come risultato $f(x)$ ". (Citato in [18, p. 112], originale in [9, pp.207-208.]).

Ora, entrambe le identità (4), (5) per noi non hanno senso, come uguaglianze tra funzioni:

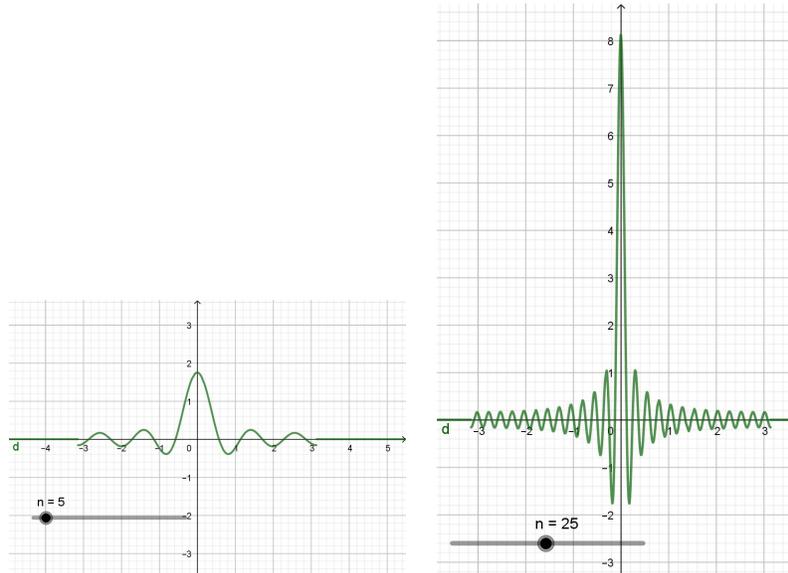
-come abbiamo già spiegato discutendo la (1), non può esistere una funzione δ che integrata contro qualsiasi f restituisca il valore di f in un certo punto: la δ dovrebbe essere "infinitamente concentrata";

-di più, la serie che compare nella (5) e dovrebbe rappresentare $\delta(x)$ non converge in nessun punto.³

Se però tracciamo il grafico delle somme parziali n -esime della serie,

$$\delta_n(x) = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^n \cos(kx) \right)$$

otteniamo delle funzioni sensate il cui grafico ha un "picco" sempre più concentrato nell'origine⁴. E' come dire che il valore dell'integrale si ottiene come limite di una successione di integrali, corrispondenti a delle funzioni δ_n che tendono a δ .



Grafici di $\delta_n(x)$ per $n = 5, n = 25$

³Si noti che questa situazione è ancora peggiore rispetto alla descrizione ingenua della delta come funzione nulla fuori dall'origine e infinita nell'origine: la serie in (5) infatti per $x = 0$ diverge a $+\infty$, ma anche per $x \neq 0$ non converge mai (in questo caso non perché diverga a ∞ ma perché è una serie oscillante).

⁴mentre fuori dall'origine ha rapide oscillazioni di segno, che fanno presagire che anche fuori dall'origine la serie non convergerà, anche se per un motivo diverso rispetto a ciò che succede nell'origine

Ed effettivamente si può dimostrare che sotto opportune ipotesi

$$\int_{-\pi}^{\pi} \delta_n(\alpha - x) f(\alpha) d\alpha \rightarrow f(x) \text{ per } n \rightarrow \infty.$$

Concludiamo questo paragrafo osservando che, sia nel trattato di Cauchy che in quello di Fourier, anche se l'argomento dell'opera è fisico-matematico (propagazione delle onde, diffusione del calore, rispettivamente), alla funzione delta non viene associato un particolare significato fisico: la delta entra in certi passaggi puramente matematici. Vedremo ora, invece, come, successivamente a questi autori, nella fisica matematica dell'800 la delta cominci ad essere investita di significati fisici.

2.2 Sorgenti concentrate e soluzioni fondamentali nella fisica matematica dell'800

Il concetto di *densità di una massa (o di una carica) puntiforme* fa la sua comparsa naturale in molti problemi fisico-matematici trattati nell'800 per mezzo di *equazioni alle derivate parziali*. Per tenere questa presentazione su un livello elementare, manterrò questa parte della discussione su linee molto generali, dando pochi dettagli in più nelle note.

Consideriamo il matematico e fisico **George Green (1793-1841)**, che nel **1828** scrive il suo *Saggio sull'applicazione dell'analisi matematica alla teoria dell'elettricità e del magnetismo* (v. [12]), una pietra miliare nella storia della fisica matematica.

Discutiamo brevemente il cosiddetto *problema fondamentale dell'elettrostatica*: assegnata, nello spazio, una certa distribuzione di cariche elettriche (puntiformi, o distribuite su fili conduttori, o su superfici conduttrici, o diffuse in un mezzo continuo tridimensionale, in ogni modo supposte *immobili*), si vuole determinare il campo elettrostatico in ogni punto dello spazio, o più precisamente si vuole determinare la funzione *potenziale elettrostatico*, $u(x, y, z)$. Per fissare le idee, occupiamoci del caso in cui la carica sia distribuita in un mezzo continuo tridimensionale, quindi si possa parlare della funzione densità (volumica) di carica, $f(x, y, z)$. Allora Green mostra che la funzione potenziale, u , risolve un'equazione differenziale

$$Lu = f$$

dove L è un certo operatore differenziale⁵, u è l'incognita e la densità f è un dato del problema.

⁵Per non fare troppi misteri, l'equazione è:

$$\Delta u = 4k\pi f$$

dove k è la costante dell'elettrostatica e Δ il laplaciano, cioè

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Ora un'idea chiave che Green introduce è la seguente. Per saper calcolare il potenziale u relativo a una qualsiasi densità f di carica è sufficiente saper calcolare il potenziale nel caso idealizzato in cui la sorgente sia una carica concentrata in un punto, quindi f sia la funzione delta, densità della carica puntiforme. Questo potenziale speciale che risolve l'equazione $Lu = \delta$ si chiama *soluzione fondamentale*⁶ e si indica con Γ . Dunque: conoscendo Γ , soluzione di $L\Gamma = \delta$, si può calcolare u soluzione di $Lu = f$, per “qualsiasi” f , con un “semplice” integrale⁷ che assegna u in funzione di Γ e f . Questo è esattamente il motivo per cui è naturale rappresentare la distribuzione di carica concentrata in un punto mediante la sua densità (delta), cioè come caso particolare di una distribuzione estesa nello spazio.

La cosa interessante è che però, a differenza della delta, la soluzione fondamentale Γ in questo caso è una funzione vera, un oggetto molto più trattabile: il potenziale elettrostatico generato nello spazio da una carica puntiforme posta nell'origine, è semplicemente la funzione

$$\Gamma(x, y, z) = \frac{c}{\sqrt{x^2 + y^2 + z^2}}$$

(per una certa costante c), che non ha nulla di misterioso. In questi problemi fisico-matematici la δ entra quindi in modo un po' indiretto: possiamo dire che *la delta è il “termine noto misterioso” di un'equazione la cui soluzione è nota e non ha niente di misterioso*. La delta rimane quindi sullo sfondo intuitivo del discorso, attraverso l'espressione verbale “carica unitaria concentrata in un punto” (la cui densità dovrebbe essere la delta), ma poi i calcoli matematici effettivi si fanno sul potenziale Γ , che è in un certo senso l'*effetto* di questa delta, ed è un oggetto noto e trattabile.

L'esempio di Green è probabilmente il primo, nell'800, in cui si considerano, sia pure in modo un po' implicito, equazioni differenziali in cui il termine noto rappresenterebbe una sorgente concentrata di tipo delta. Osserviamo che il concetto di *soluzione fondamentale* ha una portata molto più ampia, non limitata all'elettrostatica: per ogni operatore differenziale lineare L si può parlare di soluzione fondamentale, ossia soluzione di $Lu = \delta$ ⁸. Nel *Trattato sui raggi luminosi* del 1882 di **Gustav Kirchhoff (1824 – 1887)**, ritroviamo il concetto

⁶In effetti Green non considera esattamente quella che oggi chiamiamo *soluzione fondamentale*, ma quelle che oggi chiamiamo (appunto) *funzioni di Green*. La funzione di Green $G(\underline{x}, \cdot)$ di un certo dominio Ω dello spazio tridimensionale è il potenziale elettrostatico generato da una carica puntiforme unitaria posta nel punto \underline{x} di Ω , imponendo inoltre che il potenziale si annulli sul bordo di Ω (“messa a terra”). La soluzione fondamentale è quindi la funzione di Green nel caso particolare in cui il dominio Ω sia tutto lo spazio. Per il punto che mi interessa sottolineare in questa esposizione, confondere i due concetti è un peccato veniale.

⁷la formula è la seguente:

$$u(\underline{x}) = \int_{\mathbb{R}^3} \Gamma(\underline{x} - \underline{y}) f(\underline{y}) d\mathbf{y}.$$

⁸Notiamo che nessuno ci garantisce che la soluzione fondamentale di un operatore L “qualsiasi” sia sempre un oggetto altrettanto trattabile di quanto accade per l'operatore di Laplace dell'elettrostatica.

di soluzione fondamentale per l'*equazione delle onde*. Ci sono vari altri autori e lavori dell'800 o inizio '900, tra cui **Hermann von Helmholtz (1821 – 1894)** e **Lord Kelvin (1824 – 1907)**, in cui ritroviamo idee analoghe, ma non entreremo in altri dettagli. Torneremo sul concetto di soluzione fondamentale, nel contesto della teoria delle distribuzioni, nel § 5.2.1.

Un ulteriore passo di “spregiudicatezza” rispetto all'uso della delta si trova in **James Clerk Maxwell (1831 – 1879)**. Nel suo famoso *Trattato sull'elettromagnetismo*, nel 1873, dà una trattazione matematica dei *dipoli elettrici*, utilizzando una idealizzazione che si può vedere addirittura come una sorta di “derivata della funzione delta”. Un antecedente storico di quest'idea di Maxwell si ha in **Siméon-Denis Poisson (1781 – 1840)**, col suo *Trattato sul magnetismo* del 1821/22, in cui definisce gli *elementi magnetici* (dipoli magnetici) come “piccolissime parti di materia in cui i due fluidi magnetici sono presenti in parti uguali”.

2.3 Heaviside e le correnti impulsive

Consideriamo ora un autore che a fine '800 sostenne apertamente l'utilizzo di una funzione “di tipo impulsivo” (funzione delta): **Oliver Heaviside (1850 – 1925)**. Heaviside, che di professione era un ingegnere telegrafico e telefonico, si occupò principalmente di elettromagnetismo e circuiti elettrici. Il famoso trattato di Maxwell sull'elettromagnetismo fu scritto nel 1873; Heaviside lo studiò, ci lavorò sopra, ne riformulò certe parti. Ad esempio, le “equazioni di Maxwell” dell'elettromagnetismo, nella forma in cui le scriviamo oggi, sono dovute a Heaviside, non si trovano scritte così nel trattato di Maxwell. Heaviside è un personaggio con caratteristiche contrastanti. I suoi contributi all'elettromagnetismo furono apprezzati, tuttavia il mondo accademico non gradiva i suoi metodi non rigorosi. Non ostante fosse stato ammesso alla prestigiosa Royal Society, da un certo momento in poi gli fu impedito di pubblicare i suoi lavori sui Proceedings della Royal Society. Vediamo cosa c'era di non rigoroso nei metodi di Heaviside.

Per trattare i circuiti elettrici dal punto di vista matematico, negli anni 1890, Heaviside inventò un suo metodo, chiamato *calcolo operativo* o *calcolo simbolico*. Dato un circuito elettrico, in base alle leggi dell'elettrologia noi possiamo scrivere un sistema di equazioni differenziali che, una volta risolto, determina le correnti elettriche (o le tensioni) nelle varie maglie del circuito. Ora, sappiamo che gli ingegneri vorrebbero conoscere le correnti elettriche senza dover *veramente* risolvere le equazioni differenziali, e per raggiungere questo obiettivo ne inventano una più del diavolo: Heaviside fu un caposcuola di questa filosofia. Nel suo calcolo operativo lui manipolava gli operatori differenziali come se fossero delle quantità algebriche, senza dare nessuna giustificazione delle sue procedure, e miracolosamente arrivava a risultati corretti. Più di quarant'anni dopo, nel 1937, il matematico tedesco **Gustav Doetsch (1892 – 1977)** diede una sistemazione coerente e moderna alla *teoria della trasformata di Laplace* (v. [7]), che in particolare dava un senso ai risultati del calcolo di Heaviside. Ma ai suoi tempi Heaviside non fece alcun tentativo di giustificare il suo calcolo rigorosamente, e difese orgogliosamente la sua posizione pragmatica:

“Shall I refuse my dinner because I do not fully understand the process of digestion? No, not if I am satisfied with the result”. [19, p.2]

All'interno della sua teoria dei circuiti elettrici, che già di per sé aveva abbastanza motivi per dare scandalo a qualsiasi matematico o fisico rigoroso, Heaviside fece anche uso di funzioni di tipo “delta”. Vediamo come, riferendoci al suo articolo *Sugli operatori in fisica matematica*, del **1893** [14].

L'intensità di corrente in un circuito si può vedere come il tasso istantaneo di passaggio di carica elettrica attraverso una sezione del circuito. Matematicamente,

$$I(t) = \frac{dQ}{dt}.$$

Supponiamo che la carica subisca ora una variazione istantanea da 0 a Q_0 cioè sia:

$$Q(t) = Q_0 H(t)$$

dove

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

è la funzione che oggi chiamiamo appunto “gradino di Heaviside”. (Il fatto che possa esserci una variazione istantanea è un caso limite, un'idealizzazione, ma può essere comodo pensare così quando la variazione è molto rapida). L'intensità di corrente $I(t)$, che è in generale è la derivata di $Q(t)$, in questo caso dovrebbe essere la derivata della funzione gradino:

$$I(t) = \frac{d}{dt} (Q_0 H(t)) = Q_0 H'(t)$$

o, con le notazioni di Heaviside, che usava la lettera p come simbolo dell'operatore $\frac{d}{dt}$ (e poi manipolava p come se fosse una quantità algebrica),

$$Q = Q_0 \cdot pH.$$

Dal punto di vista matematico, essendo $H(t)$ costante per $t > 0$ e costante per $t < 0$, la derivata $H'(t)$ vale zero tranne che in $t = 0$, in cui non esiste. Ma dire *solo* questo è poco significativo. Per il suo significato di derivata della carica, se la carica passa da 0 a Q_0 l'integrale della corrente deve valere Q_0 (l'integrale della derivata di Q è la differenza tra i valori di Q agli estremi):

$$\int_{-\infty}^{+\infty} I(t) dt = Q_0,$$

ossia

$$\int_{-\infty}^{+\infty} Q_0 (pH) dt = Q_0$$

e quindi

$$\int_{-\infty}^{+\infty} (pH) dt = 1.$$

La funzione (pH) (cioè $H'(t)$) quindi è zero fuori dall'origine e ha integrale 1: perciò è la “funzione delta”, e rappresenta in questo caso una *corrente elettrica impulsiva*, cioè un impulso istantaneo di corrente capace di portare un incremento misurabile della carica.

La novità di Heaviside rispetto ad alcuni “pionieri della delta” citati in precedenza, che la usano un po' di nascosto, in vari punti della loro opera, senza dare un nome o introdurre un simbolo per questo oggetto, è l'esplicitezza con cui lui tratta la funzione impulsiva come un oggetto ben determinato, con un simbolo preciso, pH (“derivata del gradino”). Heaviside inoltre difende apertamente il “mostro”. Scrive infatti, commentando le formule qui sopra:

È un nonsenso? È un risultato assurdo che indica la scarsa attendibilità della matematica operativa 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 $t \geq 0$, 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 pQ coinvolge gli ordinari concetti di derivata e integrale, spinti al limite [Heaviside, [18, p.116]].

Dal punto di vista delle pretese proprietà matematiche della delta, Heaviside aggiunge a quelle che abbiamo incontrato finora il fatto che *la delta sia la derivata della funzione gradino* (che per Heaviside ne è la definizione stessa).

2.4 La delta “di Dirac”

Veniamo finalmente al personaggio che ha legato per sempre il suo nome alla funzione delta: **Paul Adrien Maurice Dirac (1902 – 1984)**, fisico britannico, premio Nobel nel 1933. La sua opera e il ruolo della delta sono legati alla nascita della *meccanica quantistica*.

Ricordiamo il contesto storico-scientifico. Tra il 1925 e il 1926 furono proposte due diverse teorie della nascente meccanica quantistica. La prima, la cosiddetta “meccanica delle matrici”, fu formulata a Göttingen da Werner Heisenberg, Max Born, Pascual Jordan. La seconda, la cosiddetta “meccanica ondulatoria”, fu formulata da Ernst (o Erwin) Schrödinger (università di Zurigo, poi Berlino), in una famosa serie di 4 paper. Senza entrare in dettagli, diciamo che dal punto di vista matematico la *meccanica delle matrici* si basava su una *matematica discreta*, con matrici infinite che rappresentano gli osservabili fisici, mentre la *meccanica ondulatoria* si basava su una *matematica del continuo*, col ruolo centrale di un'equazione differenziale, l'equazione di Schrödinger, e delle sue soluzioni, le funzioni d'onda. Poiché entrambe le teorie facevano previsioni fisiche corrette, era naturale cercare di vederle come due diversi modi, entrambi

leciti, di raccontare la stessa storia. I due racconti però erano veramente molto difficili da conciliare dal punto di vista matematico.

Nel **1930** Dirac scrisse il testo fondamentale *The principles of quantum mechanics* (v. [6]), che si rivelerà molto influente tra i fisici. In questo testo Dirac presentò una sua sintesi delle due teorie. Ma la sintesi fu compiuta senza alcun rispetto per il rigore matematico. Infatti, l'idea di Dirac⁹ è quella di mostrare l'equivalenza tra le due teorie ottenendo quella nel continuo (la meccanica ondulatoria di Schrödinger) come limite “per discretizzazione” a partire da quella nel discreto (la meccanica delle matrici di Heisenberg). In questo passaggio al limite, gli operatori che ottiene a partire da quelli della meccanica delle matrici sono degli *operatori integrali* nel continuo, ossia operatori del tipo

$$Tf(x) = \int k(x,y) f(y) dy.$$

Il problema è che questi operatori integrali dovrebbero rappresentare anche l'operatore di Schrödinger, che è un operatore differenziale. Ora, un operatore differenziale non si può rappresentare come operatore integrale! Questo punto non è una “finezza” matematica, ma un'ostruzione logica fondamentale: gli operatori differenziali sono *locali* (il valore di Tf in un punto x dipende solo dai valori di f nei punti “molto vicini” a x) mentre gli operatori integrali sono *globali* (il valore di Tf in un punto x dipende dai valori di f in tutti i punti, anche “molto lontani” da x). Dirac aggirò elegantemente il problema introducendo come possibili nuclei integrali $k(x,y)$ le sue “funzioni delta” e, peggio ancora, *derivate di funzioni delta*.

Leggiamo ora qualche stralcio dal libro di Dirac a proposito della delta:

15. The δ function. Our work in § 10 led us to consider quantities involving a certain kind of infinity. To get a precise notation for dealing with these infinities, we introduce a quantity $\delta(x)$ depending on a parameter x satisfying the conditions

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1$$

$$\delta(x) = 0 \text{ for } x \neq 0.$$

To get a picture of $\delta(x)$, take a function of the real variable x which vanishes everywhere except inside a small domain, of length ε say, surrounding the origin $x = 0$, and which is so large inside this domain that its integral over this domain is unity. (...) Then in the limit $\varepsilon \rightarrow 0$ this function will go over into $\delta(x)$.

$\delta(x)$ is not a function of x according to the usual mathematical definition of a function, which requires a function to have a definite value for each point in its domain, but is something more general,

⁹Questa illustrazione dell'idea di Dirac è una sintesi di quella che ne dà von Neumann in [30, Chap.1, §3].

which we may call an 'improper function' to show up its difference from a function defined by the usual definition. Thus $\delta(x)$ is not a quantity which can be generally used in mathematical analysis like an ordinary function, but its use must be confined to certain simple types of expression for which it is obvious that no inconsistency can arise. (v.[6, pp.58 sgg.]

Questo paragrafo 15 del libro di Dirac si può considerare l'atto di battesimo della delta "di Dirac", che da quel momento in poi sarà indissolubilmente legata al nome del fisico britannico. Si introduce un nome, un simbolo, si elenca (vedremo ora) una lista di proprietà che si utilizzeranno in seguito e se ne rivendica la liceità d'uso. In tutta l'esposizione manca qualsiasi argomento matematico di tipo dimostrativo. Da notare invece la retorica dell'esposizione: dal titolo del paragrafo, "La funzione δ " si passa in poche righe a chiamarla, in modo meno impegnativo, "una quantità $\delta(x)$ ", per arrivare nel secondo paragrafo a dichiarare candidamente " δ non è una funzione"¹⁰

Quindi Dirac procede per diverse pagine, introducendo in modo discorsivo le varie proprietà della $\delta(x)$ che utilizzerà in seguito. Elenco solo alcune delle proprietà che Dirac riporta:

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

per ogni funzione continua f (la proprietà che abbiamo visto in Cauchy e in Fourier);

$$\delta(x) = \frac{d}{dx} H(x)$$

(con H gradino di Heaviside);

$$\delta(-x) = \delta(x)$$

$$x\delta(x) = 0$$

$$\delta(ax) = a^{-1}\delta(x) \text{ per } a > 0$$

$$\int_{-\infty}^{+\infty} \delta(a-x) \delta(x-b) dx = \delta(a-b)$$

$$\frac{d}{dx} \log x = \frac{1}{x} - i\pi\delta(x).$$

Contro questo modo spregiudicato di procedere si schierò apertamente **John von Neumann (1903 – 1957)**, che nel **1932** pubblicò un altro testo fondamentale, *Matematische Grundlagen der Quantenmechanik* (Fondamenti Matematici

¹⁰adducendo, tra l'altro, una motivazione errata per questo giudizio. Dirac scrive infatti che la δ non è una funzione "perché la definizione richiede che una funzione abbia un valore ben definito in ogni punto". Ora, la sua delta *ha un valore ben definito in ogni punto* (quando Dirac scrive, la teoria della misura da almeno 25 anni ammetteva funzioni che in certi punti valessero infinito). Non è questo, quindi, il problema, piuttosto il fatto che non esista alcuna funzione che, avendo quei valori ben definiti, abbia integrale 1.

della Meccanica Quantistica) [30], che si può considerare la prima formalizzazione matematicamente rigorosa della meccanica quantistica¹¹ e, dal punto di vista matematico, l'atto di nascita della teoria degli spazi di Hilbert. Nell'introduzione della sua opera, von Neumann scrive:

“The method of Dirac (and this is overlooked today in a great part of quantum mechanical literature, because of the clarity and elegance of this theory) in no way satisfies the requirements of mathematical rigor -not even if these are reduced in a natural and proper fashion to the extent common elsewhere in theoretical physics”. [30, p. IX].

Von Neumann non rese rigoroso l'utilizzo della delta di Dirac: *lo evitò completamente*.

Termino questa sezione storica con la seguente citazione:

“La funzione δ deve aver avuto un'infanzia molto triste, poiché né i matematici né i fisici la riconoscevano come appartenente al loro dominio. Se i matematici la usavano, era come una nozione fisicamente intuitiva priva di realtà matematica (...). D'altro canto, i fisici solitamente consideravano la funzione δ , o la massa puntiforme, come una pura idealizzazione matematica che non esiste in natura”. ([18, p. 110]).

3 La nascita della teoria delle distribuzioni di Laurent Schwartz

Laurent Schwartz (1915 – 2002) è un matematico francese famoso soprattutto per la creazione della *teoria delle distribuzioni*, che gli valse nel **1950** la *medaglia Fields*, massima onorificenza in campo matematico. Il suo primo articolo su questo tema è del **1945**. Il titolo dell'articolo è un manifesto, una sintesi efficace degli scopi e dei contenuti della teoria delle distribuzioni: *Généralisation de la notion de fonction, de dérivation, de transformation de Fourier et applications mathématiques et physiques* (v. [25]). Seguirono altri suoi tre articoli negli anni precedenti il 1950, culminati in un libro onnicomprensivo, pubblicato da Schwartz in varie edizioni dal 1950 in poi (v. [22], [23], [24]).

La teoria delle distribuzioni ha effettivamente generalizzato, come dice il titolo dell'articolo citato, alcuni concetti chiave dell'analisi matematica, come quello di *funzione* e di *derivata*. In ambito più avanzato, ha generalizzato la definizione di *trasformata di Fourier*, e anche cosa si debba intendere per “soluzione di

¹¹Questo è esattamente lo scopo dell'opera, come dichiarato esplicitamente nelle prime righe della prefazione:

“The object of this book is to present the new quantum mechanics in a unified representation which, so far as it is possible and useful, is mathematically rigorous”. [30, Preface, p.VII].

un'equazione differenziale". Si può dire che sia una delle teorie che ha maggiormente segnato l'analisi matematica del 20° secolo. Come molte grandi idee, questa certamente non nasceva dal nulla, ma rappresentava piuttosto il punto di maturazione di varie idee e anticipazioni precedenti. Poiché qui ci interessa principalmente della delta, non approfondiremo la questione storica della genesi della teoria di Schwartz. Nel far questo occorrerebbe tra l'altro distinguere tra le teorie e i contributi che *oggettivamente* erano già stati sviluppati quando Schwartz elaborò la sua teoria da quelli che *soggettivamente* Schwartz conosceva e tenne presenti come fonti di ispirazione. Per qualche approfondimento su questo tema rimando al testo [18] di Lützen, in particolare Chap. 2 Part 8 e Chap. 6. Sul piano delle anticipazioni oggettive della teoria delle distribuzioni occorre citare almeno il lavoro di **Sergej Sobolev (1908 – 1989)**, che a partire dal 1936 (v. [26]), per affrontare problemi di equazioni alle derivate parziali, introdusse gli spazi di funzioni "derivabili in senso debole", che ancor oggi sono fondamentali in questo settore dell'analisi matematica.

Ora, la teoria delle distribuzioni, tra gli altri suoi meriti, ha risolto il problema della "funzione delta", cioè: all'interno di questa teoria la delta risulta un oggetto dotato di senso. Non è una funzione, ma una distribuzione. Molte delle relazioni che si sono scritte riguardo alla delta hanno senso e sono vere, *pur di reinterpretarle opportunamente*. La teoria delle distribuzioni non è quindi una sorta di "benedizione" indiscriminata su qualunque identità fosse stata scritta in precedenza riguardo questo oggetto ("adesso sappiamo che la δ ha senso, quindi vale tutto"), ma una *revisione critica* di quelle identità. In particolare, certe affermazioni che erano folli prima della teoria delle distribuzioni, rimangono folli: occorre *studiare questa teoria* quindi, almeno nei suoi concetti fondamentali, per capire come stanno le cose.

Anche se gli obiettivi oggettivamente raggiunti dalla teoria delle distribuzioni vanno ben oltre la trattazione della delta, se leggiamo l'introduzione del libro di Schwartz vediamo che la delta è ben presente nelle sue motivazioni¹², così come lo è il calcolo simbolico di Heaviside, su cui non per nulla ci siamo soffermati in precedenza. Leggiamo dunque l'inizio del libro di Schwartz (v. [24, pp.3-4 (la traduzione è mia)]):

"Sono più di 50 anni che l'ingegner Heaviside ha introdotto le sue regole di calcolo simbolico, in una memoria audace in cui calcoli matematici molto poco giustificati sono utilizzati per la soluzione di problemi fisici. Questo calcolo simbolico, o operativo, non cessa di essere sviluppato. (...). Gli ingegneri lo utilizzano sistematicamente, ognuno con la sua concezione personale, con la coscienza più o meno tranquilla; è divenuta una tecnica che "non è rigorosa ma funziona bene". Dopo l'introduzione da parte di Dirac della famosa funzione

¹²così come è ben presente nelle motivazioni del premio che gli fu conferito: se andiamo alla pagina web della Fields Medal, dove sono elencati tutti i vincitori del premio con relative motivazioni, a proposito di Schwartz leggiamo queste semplici parole: "Developed the theory of distributions, a new notion of generalized function motivated by the Dirac delta-function of theoretical physics". Si veda: <https://www.mathunion.org/imu-awards/fields-medal/fields-medals-1950>

$\delta(x)$ che sarebbe nulla per $x \neq 0$ e sarebbe infinita per $x = 0$ in modo tale che $\int_{-\infty}^{+\infty} \delta(x) dx = 1$, le formule del calcolo simbolico sono divenute ancor più inaccettabili per il rigore dei matematici. Scrivere che la funzione di Heaviside $Y(x)$ uguale a 0 per $x < 0$ e a 1 per $x \geq 0$ ha per derivata la funzione di Dirac $\delta(x)$ la cui definizione stessa è matematicamente contraddittoria, e parlare delle derivate $\delta'(x), \delta''(x) \dots$ di questa funzione privata di esistenza reale, significa oltrepassare i limiti che ci sono concessi.

Come spiegare il successo di questi metodi? Quando una tale situazione contraddittoria si presenta, è ben raro che non ne risulti una teoria matematica nuova che giustifichi, in forma modificata, il linguaggio dei fisici; e sia essa stessa sorgente importante di progresso della matematica e della fisica.

(...)

Noi abbiamo generalizzato la nozione di funzione, insieme a quella di misura, a quella di distribuzione. δ sarà una misura e non una funzione, δ' una distribuzione e non una misura”.

La teoria delle distribuzioni fa oggi parte del bagaglio standard di molti laureati in matematica. Non si può dire lo stesso per i laureati in ingegneria, che pure hanno incontrato e visto utilizzare in vari contesti (ad esempio, nella teoria della trasformata di Fourier applicata alla teoria dei segnali) concetti che solo nell’ambito delle distribuzioni hanno un senso. Come mai, settant’anni dopo il nascere di questa teoria, questa stenta ancora a entrare nel background di chi pure ne utilizza i frutti? Forse la colpa non è tutta e solo degli ingegneri (ho detto “non solo”!). Per approfondire questo punto, ho trovato illuminante la lettura del breve saggio *The impact of distributions in analysis* (contenuto in [10]), di Lars Gårding, un matematico che ha conosciuto Schwartz e assistito al nascere della teoria delle distribuzioni. Leggiamo qualche brano di quel saggio:

The book *Théorie des distributions* by Laurent Schwartz (1951), now one of the non-read classics of mathematics, has transformed many branches of analysis, and the theory is now familiar to every student who ever took an advanced mathematics course in analysis. [10, p.77].

Perché Gårding si riferisce al libro di Schwartz come a uno dei *classici non letti* della matematica? In effetti, a dispetto della grande diffusione, almeno tra i matematici, della *teoria* di Schwartz, il *libro* di Schwartz non è tra quelli che si trovino sugli scaffali di molti di noi. Tanto per cominciare, a quanto mi risulta il libro non è mai stato tradotto non solo in Italiano, ma neppure in Inglese. E’ rimasto scritto in Francese da 70 anni a questa parte. E’ questo non è un buon indizio del successo di un’opera scientifica, ai nostri tempi. Proseguiamo leggendo qualche altro passo di Gårding:

Schwartz’s book (1951) is concerned with the theory as summed up above and the precise definitions of the linear spaces involved.

In line of applications Dirac's δ -function is seen as a measure and it is proved for instance that a distribution which is non-negative for non-negative functions is a non-negative measure, that harmonic distributions are actually real analytic functions, a result which probably extends to elliptic systems with analytic coefficients, and that the solutions of hyperbolic equations of order two, as computed by Hadamard and M. Riesz, are actually distributions, and so on. These applications do not measure up to the ambitions of the theory and *one gets the impression that the author's heart is with the linear topological spaces rather than with the problems of analysis.* [10, p.79 (il corsivo nella citazione è mio).]

La frase che ho messo in corsivo forse rende almeno in parte ragione dello scarso entusiasmo con cui la teoria è stata accolta al di fuori della cerchia dei matematici. Il punto è che nella sua formulazione originaria, la teoria di Schwartz fa uso di un certo linguaggio astratto di topologia (la teoria degli *spazi vettoriali topologici*¹³ che, Gårding suggerisce, doveva interessare molto a Schwartz), un linguaggio che fuori dalla cerchia dei matematici non c'è da aspettarsi che sia conosciuto, e anche nella cerchia dei matematici non gode di simpatia generalizzata. Apriamo una parentesi. Schwartz è un matematico non facile da classificare. Da una parte mostra una profonda conoscenza e interesse per le applicazioni fisiche e ingegneristiche della matematica. D'altro canto, come matematico teorico, il suo "gusto personale" ha un sapore molto formale che risente della "corrente bourbakista", quella scuola di influenti matematici, perlopiù francesi, a cui lo stesso Schwartz apparteneva, che per qualche decennio, nella parte centrale del '900, ha sfornato libri in cui si presentava una sistematizzazione enciclopedica della matematica, in forma di teorie assiomatiche¹⁴. Questo suo gusto per le teorie formali spiega in parte l'ampio uso che Schwartz fece del linguaggio astratto della teoria degli spazi vettoriali topologici.

Aggiungo che questo linguaggio *non è intrinsecamente necessario alla teoria stessa*: esistono oggi libri di testo che presentano la teoria (nei suoi elementi essenziali, o anche in certi suoi risultati avanzati) con un investimento iniziale piuttosto contenuto in termini di astrazione matematica, e in particolare di topologia¹⁵. Tuttavia forse qualche decennio fa questi testi più "user-friendly" non esistevano, e così un certo solco è stato scavato, a separare i potenziali utilizzatori della teoria dalla teoria stessa.

Come già detto, per quanto riguarda il mondo dei matematici occorre dis-

¹³Un cenno al motivo per cui questo linguaggio topologico viene coinvolto nella teoria sarà dato nel § 4.6.

¹⁴Il lettore curioso di sapere qualcosa di più sulla "scuola Bourbaki" può consultare questi siti:

https://it.wikipedia.org/wiki/Nicolas_Bourbaki

https://fr.wikipedia.org/wiki/Nicolas_Bourbaki

<https://www.bourbaki.fr/>

¹⁵Alla fine dell'articolo ci sono alcuni consigli di letture per approfondire, che comprendono anche le indicazioni di qualche libro di testo di questo tipo.

tinguere tra l'accoglienza *del libro* di Schwartz e l'accoglienza *della teoria*. Scrive ancora Garding:

The initial success of the theory was due to an enthusiastic forward marketing and to Laurent Schwartz's conviction of the importance of distributions and the fiery lectures which he gave in several European countries after the war. I met him the first time at such a lecture in Lund in the fall of 1948. Like many in his audiences I will forever remember his powerful rendering and insistent French intonation of one of his first sentences: *Les fonctions indéfiniment différentiables et nulles en dehors d'un compact*. [10, p.79]

At the 1950 International Congress in Boston, a Fields medal went to Laurent Schwartz for the theory of distributions. This was perhaps the first time and maybe also the last time that this prize was given for a piece of soft mathematics and not for the solution of some known difficult problem. But the prize committee had shown an impressive foresight. [10, p.80]

L'autore di questo saggio non è evidentemente un grande "fan" di Schwartz, per cui pure mostra grande rispetto. La teoria delle distribuzioni viene descritta come "a piece of soft mathematics". Di solito la medaglia Fields (il massimo riconoscimento scientifico per un matematico) viene attribuita a matematici che si siano distinti non solo per contributi molto innovativi, ma per la risoluzione di problemi molto difficili, che abbiano richiesto (*anche*, se pure non solo) lo sviluppo di dimostrazioni di estrema complessità tecnica. Da questo punto di vista la teoria delle distribuzioni, al suo nascere, non ha immediatamente fornito la soluzione di molti problemi difficili che erano sul tappeto. Ha piuttosto fornito un quadro teorico naturale in cui affrontare certi problemi, da allora in poi. In questo senso, al confronto con altri lavori per cui è stata assegnata una medaglia Fields, Gårding si permette di giudicare l'opera di Schwartz "a piece of soft mathematics". Riconosce però che la giuria del premio abbia mostrato "un'impressionante preveggenza", perché la teoria di Schwartz conteneva il seme di grandi novità.

4 Le idee di base della teoria delle distribuzioni

Questa parte dell'articolo contiene un'esposizione (divulgativa, un po' semplificata, ma sostanzialmente rigorosa) delle prime idee della teoria delle distribuzioni, sufficienti a fare un po' d'ordine nelle idee attorno alla delta di Dirac.

4.1 Cos'è una funzione? La definizione puntuale

Poiché, come anticipato, l'idea di Schwartz è stata quella di generalizzare il concetto di funzione, presentiamo l'idea proprio seguendo, molto sinteticamente, il filo storico di come il concetto di funzione si sia evoluto dal 1700 al 1950.

Intorno al 1700 non esisteva una definizione universalmente accettata di funzione. Newton e i pionieri del calcolo infinitesimale in quegli anni pensavano a questo riguardo più o meno quello che pensa uno studente del prim'anno: una funzione è un'espressione analitica esplicita, costruita a partire dalle funzioni elementari (potenze, esponenziali, trigonometriche...). Questo punto di vista “ingenuo” è teorizzato da **Leonhard Euler (1707 – 1783)** a metà '700:

“Una funzione di quantità variabili è un'espressione analitica composta in modo qualunque da quelle quantità e da numeri o quantità costanti”¹⁶

Quando qualcuno dei pionieri inventava un nuovo modo di assegnare espressioni analitiche di funzioni, si allargava l'universo delle funzioni: ad esempio Newton diede diritto di cittadinanza alle serie di potenze, che divennero nuove funzioni.¹⁷

Agli inizi dell'800 cominciò il cosiddetto *movimento di rigorizzazione dell'analisi*. Negli anni 1830, sulla spinta in particolare delle ricerche legate alla convergenza delle serie trigonometriche che aveva introdotto Fourier nel suo saggio già citato *Théorie analytique de la chaleur*, del 1822, emerse la necessità di avere una definizione chiara di funzione, e fu in quel contesto che **Lejeune Dirichlet (1805 – 1859)** diede la definizione che si trova anche oggi nei libri di testo di analisi: una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$, ad esempio, è una legge che a ogni numero reale associa uno e uno solo numero reale.¹⁸ Non importa quale sia la regola, se sia una regola semplice, complicata, esplicita o implicita. Basta che concettualmente il numero $f(x)$ sia univocamente determinato per ogni x del dominio, allora possiamo dire che la funzione è ben definita. L'estrema generalità di questo concetto di “legge” costituisce il grande allargamento concettuale di questa definizione. Ad esempio: mentre i pionieri del calcolo davano per scontato che una funzione dovesse avere un grafico dato da una linea continua, Dirichlet presenta il suo esempio di funzione $f(x)$ che vale 1 se x è razionale, 0 se x è irrazionale, e risulta essere *discontinua in tutti i punti*. Un grande balzo avanti di generalità, che richiese da lì in poi molta attenzione alle *ipotesi* che facciamo. Infatti il matematico Felix Klein commentò così la situazione:

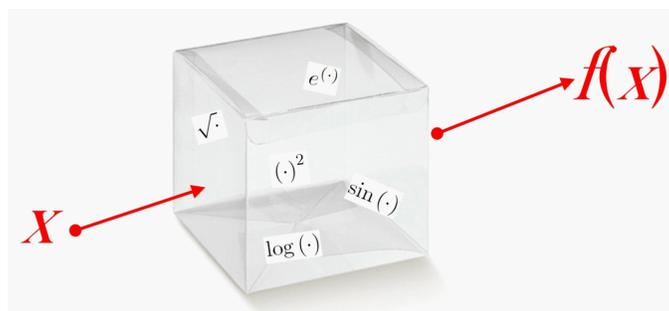
¹⁶Euler, 1748, *Introductio in analysin infinitorum*, Bousquet, Losanna. vol. 8, p. 4. E' vero che le parole “in modo qualunque” lasciano pensare una generalità totale; si osservino però anche le parole “espressione analitica”. Si può dire che la prassi matematica del tempo vincolava di fatto alla composizione di funzioni “elementari”, o comunque “accettate” a quel tempo come funzioni ammissibili.

¹⁷Altre volte invece una nuova funzione faceva il suo ingresso per un *cambiamento di punto di vista su oggetti già noti*: ad esempio, Roberval (Gilles Personne de Roberval, 1602-1675, matematico francese) e Torricelli (Evangelista Torricelli, 1608-1647) per primi tracciarono i grafici delle funzioni $\sin x$ e $\log x$, dove la cosa non banale è iniziare a pensare il seno di un angolo come una funzione dell'argomento, e tracciarne il grafico, e analogamente iniziare a vedere il logaritmo non più come un certo esponente, ma come una funzione del suo argomento.

¹⁸Una definizione di funzione come corrispondenza univoca tra due insiemi *qualsiasi*, non necessariamente “insiemi numerici” (che è la “vera” definizione moderna di funzione) si trova scritta, cent'anni dopo, in Bourbaki, 1939, [3].

“Ci accorgemmo che le funzioni, alla pari degli esseri umani, sono capaci del peggio”¹⁹

In vista dei discorsi successivi, può essere utile schematizzare la discussione precedente in questo modo pittoresco. Ciò che hanno in comune il punto di vista del '700 e dell'800 sul concetto di funzione è che la funzione viene vista come una “scatola” con un ingresso e un'uscita: si inserisce in ingresso *un numero* x e la scatola restituisce in uscita *un numero* $f(x)$ univocamente determinato. Chiamerò quest'oggetto “scatola 1.0”: *un numero in ingresso, un numero in uscita*. La *differenza* tra i due punti di vista del '700 e dell'800 è che la scatola del '700 è “trasparente”: noi pensiamo di sapere cosa ci possa essere dentro (gli “ingranaggi” sono le funzioni elementari):



La scatola dell'800, quella di Dirchlet, invece è dipinta di nero: non sappiamo quali ingranaggi possano esserci dentro la scatola: *qualsiasi* legge va bene.



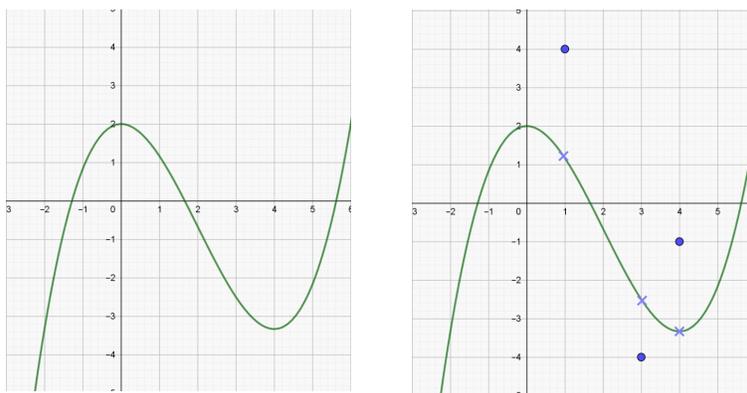
In base alla definizione dell'800, per conoscere una funzione dobbiamo sapere quanto vale *in ogni punto* (del dominio di definizione). Si parla perciò anche di *definizione puntuale* di funzione. Ufficialmente questa definizione di funzione è ancora in vigore, non è mai stata rinnegata dai matematici, ma agli inizi del '900 ha iniziato ad andare un po' stretta all'analisi matematica. Cerchiamo ora di spiegare perché.

¹⁹Felix Klein, 1849-1925, cit. in D'Amore-Matteuzzi, Dal numero alla struttura, Zanichelli, 1975, p.70.

4.2 Le funzioni nella teoria moderna dell'integrazione

Nei primi anni del '900 nasce la *teoria moderna dell'integrazione*, di **Henri Lebesgue (1875 – 1941)**, la cui importanza conquista tutti i settori dell'analisi matematica²⁰. Gli integrali diventano in un certo senso i veri protagonisti dell'analisi, perfino la definizione di soluzione di un'equazione differenziale viene via via evolvendo, con definizioni di “soluzione in senso debole” in cui si richiede che valga una certa identità tra integrali²¹. Del resto, le equazioni fondamentali della fisica matematica, come l'equazione del calore, l'equazione di continuità, vengono dedotte con opportuni procedimenti di limite a partire da equazioni di bilancio che hanno una formulazione in termini di integrali. In un certo senso, ciò che sta all'origine fisica dell'equazione differenziale, che nell'idealizzazione dell'equazione differenziale si era un po' perso nell'800, nel '900 viene posto anche a fondamento della definizione di soluzione dell'equazione differenziale stessa.

Ora, se una funzione compare dentro un integrale, non è più così cruciale sapere esattamente quanto vale la funzione *in ogni punto*: se alteriamo il valore di una funzione in due o tre punti, non cambia il suo integrale. L'integrale su (a, b) delle due funzioni in figura qui sotto, ad esempio, è lo stesso (e questo è vero per qualsiasi intervallo (a, b)):



La teoria di Lebesgue ha introdotto il concetto di funzioni *uguali tra loro quasi ovunque*; in questa teoria due funzioni uguali tra loro quasi ovunque vengono identificate, trattate come se fossero la stessa funzione. Se ad esempio due funzioni $f, g : [a, b] \rightarrow \mathbb{R}$ sono uguali in tutti i punti di $[a, b]$ tranne un numero finito di punti, possiamo dire che $f = g$ quasi ovunque in $[a, b]$. La definizione precisa di questa nozione è più sottile e permette che le due funzioni siano

²⁰Il primo lavoro fondamentale in questa direzione è la *tesi di dottorato* di Lebesgue, pubblicata nel 1902 sugli Annali di Matematica Pura e Applicata col titolo *Intégrale, Longueur, Aire*. Nel 1904 sarà pubblicata una monografia di Lebesgue dedicata all'integrazione. Questo è solo il punto di avvio di ricerche continuate nei decenni successivi, da Lebesgue e molti altri matematici.

²¹Questo ha a che fare con l'accenno fatto in precedenza alle teorie sviluppate da Sobolev a partire dal 1936, v. §3.

diverse anche in certi insiemi infiniti di punti, ma per il seguito del discorso la precedente idea intuitiva può bastare²².

Abbiamo fatto un passo importante: la teoria moderna dell'integrazione, dal '900 in poi, identifica due funzioni quando sono *uguali quasi ovunque*²³. E' come se una persona un po' miope guardasse il grafico di una funzione tracciato sulla lavagna dalla distanza di qualche metro: se qualcuno altera il valore della funzione in pochi punti, lui non se ne accorge neppure. Ora Dirichlet si rivolta nella tomba, perché se una funzione è definita così, "a meno di uguaglianza quasi ovunque", non si può più giurare su quanto valga la funzione in *nessun* punto specifico.

4.3 Le funzioni come funzionali

Proseguiamo nell'idea di dare un ruolo centrale al concetto di integrale, in analisi. Abbiamo visto che se alteriamo una funzione in pochi punti, il valore del suo integrale non cambia. D'altro canto, è chiaro che l'integrale di una funzione f , da solo, non determina certo la funzione: tante funzioni diverse possono avere lo stesso integrale.

L'idea allora è considerare non solo quanto fa l'integrale di f ma anche quanto fa l'integrale

$$\int_{\mathbb{R}} f(x) \phi(x) dx$$

al variare di ϕ in tutti i modi possibili in una certa classe di funzioni "di prova". In altre parole, invece di descrivere una funzione f chiedendoci: "Per ogni $x \in \mathbb{R}$, quando vale $f(x)$?" ora ci chiediamo: "Per ogni funzione ϕ di prova (dovremo precisare cosa significa), quando vale $\int_{\mathbb{R}} f(x) \phi(x) dx$?"

Stiamo cambiando la *domanda fondamentale* da fare alle funzioni. Data una f , invece di chiederle quanto vale in un certo punto x , ora le chiediamo quanto vale il suo integrale contro una certa funzione ϕ . Stiamo cambiando il tipo di "scatola" con cui descriveremo le funzioni.

²²Per il lettore interessato, diamo comunque la definizione esatta:

diciamo che due funzioni $f, g : I \rightarrow \mathbb{R}$ ($I \subset \mathbb{R}$) sono uguali tra loro quasi ovunque in I se l'insieme E dei punti $x \in I$ tali che $f(x) \neq g(x)$ ha misura di Lebesgue nulla, cioè soddisfa la seguente condizione:

per ogni $\varepsilon > 0$ esiste una successione di intervalli (a_n, b_n) ($n = 1, 2, 3, \dots$) tale che:

$$E \subset \bigcup_{n=1}^{\infty} (a_n, b_n)$$

$$\sum_{n=1}^{\infty} |b_n - a_n| < \varepsilon.$$

Si dimostra che, ad esempio, ogni insieme *numerabile* ha misura nulla. Poiché l'insieme dei razionali è numerabile, da questo segue che la "scandalosa" funzione di Dirichlet che vale 1 nei razionali e 0 negli irrazionali, dal punto di vista della teoria di Lebesgue è uguale a zero quasi ovunque, quindi è identificabile con la funzione identicamente nulla, che non ha nulla di scandaloso.

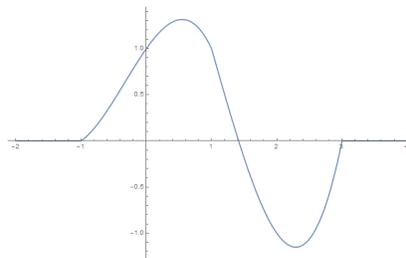
²³Il modo formale in cui ci si esprime è il seguente: invece di considerare le vecchie funzioni, si considerano *classi d'equivalenza di funzioni*. Ogni classe d'equivalenza consiste, poniamo, di una certa funzione f e tutte quelle che sono uguali quasi ovunque a f .

Fissiamo le idee. Anzitutto, quali sono le funzioni f che vogliamo descrivere in questo modo diverso dal solito? Consideriamo le funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ che siano (almeno) integrabili su ogni intervallo limitato $[a, b]$. Si dicono “funzioni localmente integrabili”. Ad esempio, sono funzioni di questo tipo le funzioni continue, ma anche certe funzioni discontinue come la funzione gradino di Heaviside, e molte altre.

Ora diciamo quale è la classe delle “funzioni di prova”, le ϕ contro cui le nostre funzioni f dovranno essere integrate (e dovranno saperci dire quanto vale l’integrale) Per essere sicuri che abbia senso l’integrale

$$\int_{\mathbb{R}} f(x) \phi(x) dx$$

(dove f , ricordiamo, è una funzione integrabile su ogni intervallo limitato -ma potrebbe non essere integrabile su tutto \mathbb{R} , si pensi a $f(x) = e^x$) è bene che le ϕ siano funzioni continue che valgono zero fuori da un intervallo limitato, così:

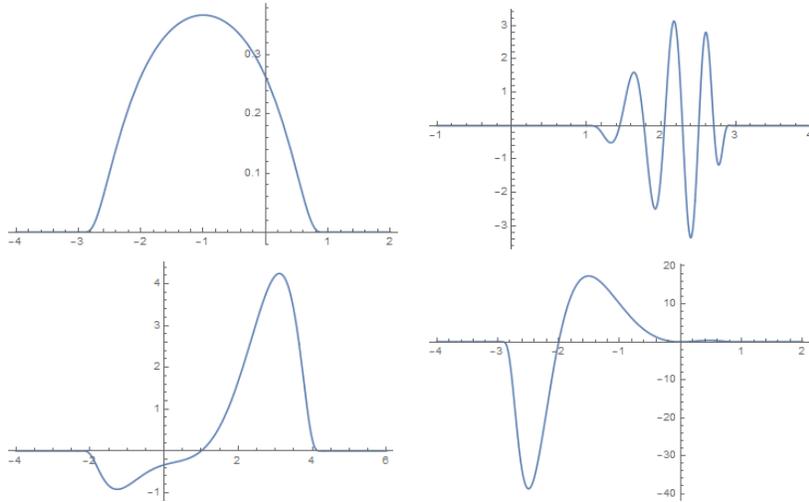


In vista degli sviluppi successivi della teoria, è utile che siano anche funzioni *regolari*, cioè derivabili tutte le volte che ci servirà. Allora “esageriamo” e diamo questa definizione molto esigente:

Definizione 1 *Lo spazio delle “funzioni test” è il seguente:*

$$\mathcal{D}(\mathbb{R}) = \{ \phi : \mathbb{R} \rightarrow \mathbb{R} : \phi \text{ è derivabile infinite volte ed è nulla fuori da un intervallo limitato} \}.$$

Per non equivocare la richiesta di annullamento precisiamo: ogni funzione test si annulla fuori da un intervallo che può essere diverso da funzione a funzione. Qualche esempio di funzione test si trova nei grafici qui di seguito:



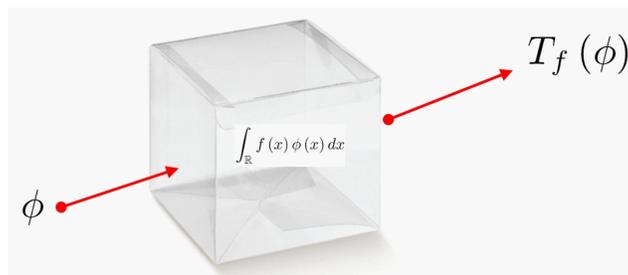
Ora: ad ogni funzione f localmente integrabile associamo il *funzionale* T_f

$$T_f : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$$

$$T_f : \phi \mapsto \int_{\mathbb{R}} f(x) \phi(x) dx.$$

La parola “funzionale” indica una legge che ad ogni funzione (di un certo tipo) associa un numero²⁴.

Questo nuovo modo di rappresentare una funzione f corrisponde ad un diverso standard di “scatola”. La chiamerò “scatola 2.0”:



La scatola ha sempre un ingresso e un’uscita, ma ora nell’ingresso non inseriamo un numero x ma una *funzione test* ϕ ; l’uscita invece è ancora un numero.

²⁴Naturalmente, in base alla definizione astratta di funzione, anche un funzionale è una particolare funzione, ma usare due termini diversi aiuta a non confondersi: di solito chiamiamo funzione una legge che associa a ogni numero un numero, e funzionale una legge che associa ad ogni funzione un numero.

Questa scatola è *perfettamente trasparente*: all'interno agisce sempre lo stesso tipo di ingranaggio, la stessa regola: "Prendi f , calcola il suo integrale contro la funzione test ϕ , e restituisci in uscita il risultato numerico dell'integrale".

La prima volta che si incontra questa scatola come modo di rappresentare le vecchie funzioni si rimane un po' perplessi, per almeno due motivi:

1. Sembra che si sia trovato un modo più complicato per descrivere una cosa semplice.

2. Siamo sicuri che questo modo di rappresentare le funzioni sia adeguato? Cioè, conoscendo i valori $T_f(\phi)$ per ogni ϕ conosciamo davvero la funzione, oppure abbiamo perso informazioni, rispetto a sapere quanto valeva f in ogni punto (o in "quasi ogni" punto)?

Sulla prima perplessità, lasciamo per ora le cose in sospeso: inutile negare che sia un po' complicato, diciamo solo che saranno i successi della teoria "a valle" a convincerci che ne valeva la pena. Veniamo alla seconda perplessità. Anzitutto, se alteriamo f in un numero finito di punti (o più in generale, se sostituiamo f con una funzione g uguale a f quasi ovunque), T_f non cambia perché ciascuno degli integrali $\int f\phi$ non cambia. Questo modo di rappresentare le funzioni è quindi più adeguato di quello di Dirichlet per rappresentare le funzioni così come sono trattate dalla teoria dell'integrazione del '900: la funzione f era definita "a meno di uguaglianza quasi ovunque", quindi $f(x)$ aveva smesso di avere un valore univoco; ma il funzionale T_f ha, per ogni funzione test ϕ , un valore che rimane univocamente determinato anche alterando la funzione f in qualche punto. *Un nuovo tipo di univocità ha sostituito la vecchia, che non era più adeguata.* In sintesi:

$$(f = g \text{ quasi ovunque}) \Rightarrow (T_f(\phi) = T_g(\phi) \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R})).$$

La cosa interessante è che vale anche il viceversa. La nostra perplessità era: questa descrizione alternativa di f riesce davvero a cogliere tutte le informazioni che ci sono in f ? Non è *troppo poco* sapere quanto vale l'integrale di f contro una funzione test qualsiasi? No, non è troppo poco e ce lo garantisce un teorema:

Teorema 2 *Se $f, g : \mathbb{R} \rightarrow \mathbb{R}$ sono due funzioni localmente integrabili e*

$$\int_{\mathbb{R}} f(x) \phi(x) dx = \int_{\mathbb{R}} g(x) \phi(x) dx \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}),$$

allora

$$f(x) = g(x) \text{ quasi ovunque.}$$

Complessivamente possiamo quindi affermare che:

$$T_f = T_g \Leftrightarrow f = g \text{ quasi ovunque.}$$

Possiamo quindi concludere che questa "scatola 2.0" dà una *rappresentazione fedele* delle funzioni, così come le intende la teoria dell'integrazione del '900: una funzione f e tutte quelle ad essa uguali quasi ovunque sono rappresentate da *un'unica* scatola 2.0, un unico funzionale, che contiene in sé tutta l'informazione contenuta nella funzione f . Inoltre questa scatola è di nuovo un box che associa un'uscita *univoca* $T_f(\phi)$ ad ogni ingresso ϕ , diversamente dalle funzioni della teoria di Lebesgue.

4.4 Nuovi funzionali. Le distribuzioni

Abbiamo fatto un passo avanti importante, eppure se ora ci fermassimo qui avremmo solo riformulato in un altro modo il concetto di funzione, senza generalizzare questo concetto, senza ampliare l'universo delle funzioni. Ma adesso siamo pronti per compiere il grande balzo in avanti, che consiste nel considerare *altri* funzionali, *che non si ottengono col procedimento visto a partire da funzioni*. Così si allargherà effettivamente l'universo dei nostri oggetti. L'idea di distribuzione di Schwartz consiste “semplicemente”, a questo punto, nel prendere la scatola 2.0 che abbiamo appena descritto e *dipingerla di nero*. Cioè dire:

Definizione 3 Chiameremo distribuzione, in \mathbb{R} , ogni funzionale

$$T : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$$

che sia lineare e continuo [in un senso che andrà precisato]. Per dire che T è una distribuzione in \mathbb{R} scriveremo

$$T \in \mathcal{D}'(\mathbb{R}),$$

dove il simbolo $\mathcal{D}'(\mathbb{R})$ indica appunto lo spazio di tutti i funzionali lineari continui su $\mathcal{D}(\mathbb{R})$.



Lasciamo un momento in sospenso l'onere di dare un significato preciso ai due termini *lineare* e *continuo* (lo faremo nel § 4.6, per il lettore interessato). Per chi vuole capire l'idea generale di *distribuzione*, questo è tutto sommato un aspetto tecnico. Il concetto chiave è capire che una distribuzione è una “qualsiasi” legge T che ad ogni funzione test ϕ associa un numero $T(\phi)$. Il valore numerico $T(\phi)$ come è stato ottenuto? Dipende! Non sempre con lo stesso procedimento; in particolare, non necessariamente è stato ottenuto calcolando un integrale. Ad esempio, sono distribuzioni i seguenti funzionali:

$$\begin{aligned} T : \mathcal{D} &\rightarrow \mathbb{R} \\ T : \phi &\mapsto \phi'(3) \end{aligned}$$

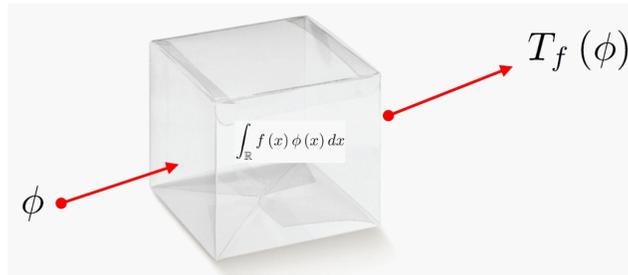
oppure:

$$T : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$$

$$T : \phi \mapsto \phi(0).$$

Si tratta di leggi che associano ad ogni funzione test ϕ un numero ben determinato (e rispettano anche le richieste di linearità e continuità, anche se non le abbiamo ancora spiegate). Il secondo funzionale che abbiamo scritto è esattamente la delta di Dirac. Vediamo di capire perché. Per più di cent'anni si è preteso che potesse esistere una funzione $\delta(x)$ tale che, tra le altre cose,

$$\int_{\mathbb{R}} f(x) \phi(x) dx = \phi(0) \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}).$$



L'idea di Schwartz (“dipingere di nero la scatola 2.0”) è questa: invece di pensare che la δ sia una funzione che integrata contro ϕ restituisce $\phi(0)$, pensiamo semplicemente che la δ sia un funzionale che ad ogni ϕ associa $\phi(0)$ (senza bisogno di calcolare nessun integrale!). Quindi la δ è una distribuzione che non si identifica con nessuna funzione.

Se invece f è una “vera” funzione (localmente integrabile), possiamo sempre costruire al vecchio modo un funzionale T_f tale che

$$T_f(\phi) = \int_{\mathbb{R}} f(x) \phi(x) dx \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}).$$

Quindi *ogni funzione* (localmente integrabile) è (identificabile con) *una particolare distribuzione*. Ma, accanto a queste distribuzioni, ce ne sono altre, come la delta di Dirac o il primo esempio qui sopra ($T : \phi \mapsto \phi'(3)$) che non sono identificabili con nessuna funzione perché, come detto ormai molte volte, non esiste nessuna funzione che integrata contro qualsiasi ϕ test restituisca sempre $\phi(0)$ (o $\phi'(3)$).

Abbiamo quindi effettivamente generalizzato l'universo delle funzioni, ammettendo anche oggetti più generali, come la delta e tanti altri.

Fissiamo qualche aspetto di linguaggio sulle distribuzioni. Se T è una distribuzione e ϕ una funzione test, il valore $T(\phi)$ si indica di solito col simbolo

$$\langle T, \phi \rangle$$

(che si legge *crochet*, alla francese, o *pairing*, all'inglese, o dualità, o azione del funzionale T su ϕ). Ad esempio, si scrive

$$\langle \delta, \phi \rangle = \phi(0) \text{ e NON } \int_{\mathbb{R}} \delta(x) \phi(x) dx = \phi(0).$$

Dobbiamo pensare che una distribuzione T (ad esempio la δ) non è definita “nei punti x ” ma è solo definito il suo valore “sulle funzioni test ϕ ”. Dobbiamo “digerire” il fatto che abbiamo *allargato* l'universo delle funzioni a un universo di oggetti più generali, le distribuzioni. La delta, ad esempio, *non è un nuovo tipo di funzione*. Semplicemente *non è una funzione*, ma una distribuzione, cioè un funzionale lineare continuo sullo spazio delle funzioni test. Possiamo finalmente dire cos'è, e non solo cosa non è.

Se f è una funzione (localmente integrabile), allora f si identifica con la distribuzione T_f che agisce così:

$$\langle T_f, \phi \rangle = \int_{\mathbb{R}} f(x) \phi(x) dx,$$

e in questo caso non c'è niente di male a scrivere l'integrale (ma è l'integrale di f , non l'integrale di T_f).

4.5 Derivata di una distribuzione

Sulle distribuzioni si possono fare molte operazioni. Ad esempio, si può definire la *derivata di una distribuzione*.

L'idea è: se f è una vera funzione derivabile, con derivata f' continua, sia f che f' si possono identificare con opportune distribuzioni.

$$\begin{aligned} \langle T_f, \phi \rangle &= \int_{\mathbb{R}} f(x) \phi(x) dx \\ \langle T_{f'}, \phi \rangle &= \int_{\mathbb{R}} f'(x) \phi(x) dx \end{aligned}$$

Ma noi, d'altro canto, vorremmo definire un'operazione di derivata di una distribuzione in modo che *la nuova definizione sia coerente con la vecchia*, cioè in modo che

$$(T_f)' = T_{f'}.$$

L'uguaglianza scritta qui sopra significa: se esiste la derivata di f in senso classico, vogliamo che la distribuzione che si ottiene derivando T_f (secondo una regola che non conosciamo ancora, e stiamo cercando di inventare) coincida con la distribuzione associata alla derivata “classica” f' . D'altro canto si ha:

$$\langle T_{f'}, \phi \rangle = \int_{\mathbb{R}} f'(x) \phi(x) dx$$

integrando per parti, poiché ϕ si annulla fuori da un intervallo limitato

$$= - \int_{\mathbb{R}} f(x) \phi'(x) dx = - \langle T_f, \phi' \rangle.$$

Per quelle particolari distribuzioni T_f che “provengono” da una funzione derivabile vogliamo quindi che valga

$$\langle T_{f'}, \phi \rangle = - \langle T_f, \phi' \rangle.$$

Allora per una distribuzione T qualsiasi prenderemo questa come definizione:

Definizione 4 Per ogni $T \in \mathcal{D}'(\mathbb{R})$, poniamo

$$\langle T', \phi \rangle = - \langle T, \phi' \rangle \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}).$$

La definizione appena data è l'unica possibile se vogliamo che il nuovo concetto di derivata contenga come caso particolare quello classico. Una delle cose che rendono “semplice e naturale”, da un certo punto di vista, la teoria delle distribuzioni, è che quello che abbiamo appena visto accadere nel caso del concetto di derivata è una regola generale: ogni nuova operazione (derivata di una distribuzione, convoluzione di due distribuzioni, trasformata di Fourier di una distribuzione...) viene definita in modo che, se applicata a quelle particolari distribuzioni che sono funzioni per le quali l'operazione in questione ha senso, si ottenga l'operazione che conosciamo già. Il bello è che questo criterio è *sufficiente* a determinare univocamente la definizione, che quindi non richiede particolare inventiva: si dà la definizione nell'unico modo in cui è ragionevole darla.

Da questa definizione si capisce perché è stato “furbo” scegliere come spazio di funzioni test uno spazio di funzioni derivabili *infinite volte*: se $\phi \in \mathcal{D}(\mathbb{R})$ anche $\phi' \in \mathcal{D}(\mathbb{R})$ e quindi per qualsiasi distribuzione T e qualsiasi $\phi \in \mathcal{D}(\mathbb{R})$ ha senso considerare $\langle T, \phi' \rangle$: lo spazio delle funzioni test è *stabile rispetto all'operazione di derivazione*. L'infinito in matematica dà tanti problemi ma alle volte ne risolve anche qualcuno.

Naturalmente occorrerebbe dimostrare che il funzionale T' non solo è ben definito su $\mathcal{D}(\mathbb{R})$ ma è anche lineare e continuo, quindi è effettivamente una distribuzione. Questo fa parte del tecnicismo in cui non entriamo (ma rimandiamo alla prossima sezione per qualche dettaglio in più).

Esempio 5 Facciamo l'esempio che ci interessa di più: calcoliamo la derivata del gradino di Heaviside

$$H(x) = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x < 0. \end{cases}$$

Poiché H è localmente integrabile, possiamo vederla come distribuzione, e quindi derivarla. Si ha:

$$\begin{aligned} \langle (T_H)', \phi \rangle &= - \langle T_H, \phi' \rangle = (\text{poiché } H \text{ è una funzione localmente integrabile}) \\ &= - \int_{\mathbb{R}} H(x) \phi'(x) dx = - \int_0^{+\infty} \phi'(x) dx = -(\phi(\infty) - \phi(0)) = \phi(0). \end{aligned}$$

Quindi

$$\langle (T_H)', \phi \rangle = \phi(0) = \langle \delta, \phi \rangle$$

e questo significa che

$$(T_H)' = \delta$$

ossia: la derivata distribuzionale della funzione di Heaviside è la delta di Dirac. (Si osservi che due distribuzioni sono uguali se agiscono allo stesso modo su ogni ϕ test).

Siamo partiti da un oggetto, il gradino, che è una distribuzione (ed è anche una funzione, ma come funzione non è derivabile) e derivandolo abbiamo ottenuto ancora una distribuzione (che però non è più una funzione): lo spazio delle distribuzioni è stabile rispetto all'operazione di derivazione (così come lo è lo spazio delle funzioni test, mentre non lo è lo spazio di *tutte* le funzioni localmente integrabili).

La cosa notevole è che la derivata di una distribuzione è ancora una distribuzione, quindi in particolare *ogni distribuzione è infinitamente derivabile*.

In particolare, ogni funzione localmente integrabile (anche discontinua, con punti angolosi, ecc.), *nel senso delle distribuzioni* è derivabile infinite volte. Solo che le sue derivate non sono più funzioni, in generale, ma distribuzioni.

Esempio 6

$$\langle \delta', \phi \rangle = -\langle \delta, \phi' \rangle = -\phi'(0) \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}).$$

La derivata della distribuzione delta di Dirac è la distribuzione che ad ogni ϕ associa $-\phi'(0)$. Si osservi che δ' , così come δ , agisce sulle test in modo locale, non globale.

Notiamo i corsi e ricorsi della storia. Per i pionieri del calcolo, intorno al 1700, “tutto era derivabile”. Il rigore dell’800 ci ha consegnato funzioni discontinue in tutti i punti (Dirichlet), funzioni continue e mai derivabili (Weierstrass), e ogni sorta di patologia. Gli inizi del ’900, con la teoria della misura, hanno approfondito l’abisso delle mostruosità²⁵. Con la teoria delle distribuzioni, nel 1950, tutto torna ad essere derivabile infinite volte, pur di cambiare prospettiva.

4.6 Le distribuzioni come funzionali lineari continui*

In questo paragrafo spieghiamo, per il lettore interessato, le due parole chiave che completano la definizione di distribuzione, ossia cosa significa in questo contesto *funzionale lineare e continuo*. Questo paragrafo non è necessario per cogliere il messaggio principale del discorso generale fatto in questa sezione, anche se ovviamente è indispensabile per conoscere l’esatta definizione matematica di distribuzione.

La definizione di *linearità* è la solita che si incontra nei corsi universitari del primo anno: un funzionale

$$T : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$$

²⁵ con esempi di funzioni “non misurabili”, definite con procedimenti non costruttivi, oppure esempi di funzioni “singolari”, che sono continue in ogni punto, derivabili quasi ovunque, ma tali che l’integrale della derivata non restituisce l’incremento della funzione.

è lineare se per ogni $\phi_1, \phi_2 \in \mathcal{D}(\mathbb{R})$ e $\lambda_1, \lambda_2 \in \mathbb{R}$ si ha

$$T(\lambda_1\phi_1 + \lambda_2\phi_2) = \lambda_1T(\phi_1) + \lambda_2T(\phi_2)$$

Esempi di funzionali lineari su $\mathcal{D}(\mathbb{R})$ sono:

$$T(\phi) = \int_{\mathbb{R}} \phi(x) e^x dx$$

$$T(\phi) = 3\phi(0) - 5\phi'(1).$$

Esempi di funzionali non lineari su $\mathcal{D}(\mathbb{R})$ sono:

$$T(\phi) = \int_{\mathbb{R}} \phi(x)^2 dx$$

$$T(\phi) = 2\phi(0) + 1.$$

La definizione di continuità di un funzionale lineare richiede invece preliminarmente la definizione di una nozione di convergenza per le successioni di funzioni test. L'aspetto più tecnico della teoria (nei suoi passi iniziali) è proprio questo.

Definizione 7 Sia $\{\phi_n\}_{n=1}^{\infty}$ una successione di funzioni test (cioè appartenenti a $\mathcal{D}(\mathbb{R})$). Si dice che

$$\phi_n \rightarrow 0 \text{ in } \mathcal{D}(\mathbb{R})$$

se:

- 1) esiste un intervallo $[a, b]$ fuori dal quale tutte le $\phi_n(x)$ sono nulle;
- 2)

$$\max_{x \in [a, b]} |\phi_n(x)| \rightarrow 0 \text{ per } n \rightarrow \infty;$$

- 3) la stessa condizione vale per le derivate di ogni ordine di ϕ_n : per ogni $k = 1, 2, 3, \dots$ si ha

$$\max_{x \in [a, b]} \left| \frac{d^k \phi_n}{dx^k}(x) \right| \rightarrow 0 \text{ per } n \rightarrow \infty.$$

A questo punto si può completare la definizione di distribuzione:

Definizione 8 Sia $T : \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R}$ un funzionale lineare. Si dice che T è continuo se per ogni successione di funzioni test $\{\phi_n\}_{n=1}^{\infty}$ si ha:

$$\text{se } \phi_n \rightarrow 0 \text{ in } \mathcal{D}(\mathbb{R}) \text{ allora } \langle T, \phi_n \rangle \rightarrow 0.$$

Si noti che $\langle T, \phi_n \rangle$ è una successione di numeri reali, quindi la convergenza a zero di questa successione ha il solito significato.

Definizione 9 Lo spazio delle distribuzioni su \mathbb{R} è per definizione l'insieme, indicato con $\mathcal{D}'(\mathbb{R})$, di tutti i funzionali lineari continui su $\mathcal{D}(\mathbb{R})$. Se T è una distribuzione e ϕ una funzione test, il valore $T(\phi)$ si indica col simbolo $\langle T, \phi \rangle$.

Con ciò la definizione rigorosa dello spazio delle distribuzioni su \mathbb{R} è completa. Il tutto si può poi generalizzare allo spazio delle distribuzioni in più variabili e su un dominio diverso dallo spazio intero, cioè $\mathcal{D}'(\Omega)$ con Ω aperto di \mathbb{R}^n , ma qui non ci interessa procedere in questa direzione.

Illustriamo invece su qualche semplice esempio come si utilizzano le definizioni precedenti.

Esempio 10 a. Verifichiamo che la delta δ è una distribuzione:

$$\begin{aligned}\delta &: \mathcal{D}(\mathbb{R}) \rightarrow \mathbb{R} \\ \delta &: \phi \mapsto \phi(0).\end{aligned}$$

La linearità è ovvia; per provare la continuità, sia $\phi_n \rightarrow 0$ in $\mathcal{D}(\mathbb{R})$, allora

$$\langle \delta, \phi_n \rangle = \phi_n(0) \rightarrow 0$$

perché per definizione di convergenza in $\mathcal{D}(\mathbb{R})$ si ha (non solo, ma in questo caso ci basta)

$$\max |\phi_n(x)| \rightarrow 0 \text{ per } n \rightarrow \infty.$$

b. Verifichiamo che il funzionale

$$\langle T, \phi \rangle = \phi'(3)$$

è una distribuzione. Di nuovo, la linearità è ovvia, e se $\phi_n \rightarrow 0$ in $\mathcal{D}(\mathbb{R})$, allora

$$\langle T, \phi_n \rangle = \phi_n'(3) \rightarrow 0$$

perché la definizione di convergenza in $\mathcal{D}(\mathbb{R})$ implica anche che sia

$$\max |\phi_n'(x)| \rightarrow 0 \text{ per } n \rightarrow \infty.$$

c. Infine, verifichiamo che le funzioni localmente integrabili sono distribuzioni. Sia f localmente integrabile e sia

$$\langle T_f, \phi \rangle = \int_{\mathbb{R}} f(x) \phi(x) dx.$$

La linearità di T_f è ovvia. Sia $\phi_n \rightarrow 0$ in $\mathcal{D}(\mathbb{R})$, allora per definizione di convergenza in $\mathcal{D}(\mathbb{R})$ esiste $[a, b]$ fuori dal quale tutte le ϕ_n si annullano; d'altro canto f è localmente integrabile, quindi il suo integrale su $[a, b]$ è finito; infine, ancora per definizione di convergenza in $\mathcal{D}(\mathbb{R})$ si ha

$$\max |\phi_n(x)| \rightarrow 0 \text{ per } n \rightarrow \infty.$$

Usando di questi tre fatti possiamo scrivere:

$$\begin{aligned}|\langle T_f, \phi_n \rangle| &= \left| \int_{\mathbb{R}} f(x) \phi_n(x) dx \right| = \left| \int_a^b f(x) \phi_n(x) dx \right| \\ &\leq \int_a^b |f(x) \phi_n(x)| dx \leq \max |\phi_n| \int_a^b |f(x)| dx \\ &= c \cdot \max |\phi_n| \rightarrow 0.\end{aligned}$$

Dunque T_f è continuo.

Con ragionamenti analoghi si può dimostrare che se T è una distribuzione anche T' lo è, come affermato nel paragrafo precedente.

Osservazione 11 (Convergenza e topologia) *Per il lettore arrivato fin qui, possiamo ora spiegare un po' meglio l'affermazione fatta in precedenza (v. § 3) sul fatto che la teoria delle distribuzioni, nella formulazione originale di Schwartz, facesse un uso abbastanza pesante, e non del tutto necessario, del linguaggio astratto degli spazi vettoriali topologici.*

Riprendiamo la definizione che abbiamo dato di convergenza di una successione ϕ_n di funzioni test. Dare quella definizione, e la successiva definizione di distribuzione, come abbiamo fatto, è perfettamente rigoroso (e, come si è visto, non ha richiesto particolare astrazione). Questo è un percorso didattico “normale”, oggi. Allo stesso tempo, questo percorso logico, dal punto di vista di un matematico teorico con un gusto un po' formale (come certamente era Schwartz), non è “estetivamente soddisfacente”. Quando in matematica si dice che una successione di oggetti di qualche tipo converge a qualcosa, “di solito” questo è l'applicazione a un caso concreto di una soggiacente struttura di norma, o di distanza, o di topologia, rispetto alla quale si ha quella convergenza. Ad esempio, la convergenza uniforme di una successione di funzioni è la convergenza “nella norma C^0 ”, la convergenza in media quadratica (che si usa ad esempio in analisi di Fourier) è la convergenza “nella norma L^2 ”, e così via. Ora, la definizione di convergenza di funzioni test che abbiamo dato esprime una convergenza rispetto a qualche norma o qualche distanza? La risposta è no. Esprime una convergenza rispetto a qualche topologia? La risposta è: sì, rispetto a una topologia piuttosto complicata da descrivere. Schwartz non volle rinunciare a descrivere quella topologia, inquadrandola in una teoria generale astratta degli spazi vettoriali topologici²⁶, per poi far discendere quella definizione di convergenza in $\mathcal{D}(\mathbb{R})$ come applicazione “naturale” della teoria generale. Così facendo ha costruito una teoria molto elegante e raffinata ma, diciamolo francamente, si è anche perso un po' di lettori.

5 Epilogo. Cosa abbiamo imparato?

In quest'ultima sezione vorrei cercare di raccogliere alcuni elementi che la vicenda narrata ci può insegnare.

Per prima cosa (§ 5.1) credo sia utile, dopo che nella sezione precedente abbiamo discusso ciò che effettivamente la delta è (e non solo ciò che non è), riesaminare con qualche dettaglio in più cosa ci sia che non va in certi approcci “ingenui” alla delta che si sono descritti in precedenza. Non faccio questo per inferire su posizioni già ampiamente criticate in precedenza. Piuttosto, credo che anche da questi errori ci sia qualcosa da imparare, in quanto le stesse misconcezioni spuntano anche in altri contesti, negli approcci “pragmatici” alla matematica.

²⁶Un riferimento autorevole per la teoria degli spazi vettoriali topologici è la monografia [29].

In secondo luogo, vorrei affrontare la domanda: valeva davvero la pena di mettere in atto quel percorso di ripensamento profondo di concetti di base dell'analisi, come funzione, derivata, per risolvere i problemi visti? Tutta quella fatica era necessaria? Non potevamo tenerci le "curve a campana" sempre più appuntite a cui si riferisce Dirac, era poi così grave?

A questo proposito, ricordiamo la frase che Schwartz scrive nell'introduzione della sua opera sulle distribuzioni, nel passo già citato nel § 3:

“Quando una tale situazione contraddittoria si presenta, è ben raro che non ne risulti una teoria matematica nuova che giustifichi, in forma modificata, il linguaggio dei fisici; e sia essa stessa sorgente importante di progresso della matematica e della fisica”. [24, p.4]

Schwartz dunque riteneva (e certamente si augurava) che la sua teoria sarebbe stata “essa stessa sorgente importante di progresso della matematica e della fisica”, e che valesse la pena lavorare per elaborare una teoria “che giustifichi il linguaggio dei fisici”.

Nel § 5.2 daremo qualche breve cenno (restando molto in superficie, perché un discorso più approfondito diventerebbe molto difficile) su alcuni elementi di progresso che la teoria delle distribuzioni ha portato nello sviluppo della matematica teorica.

Questo lascia aperta l'ultima domanda, e cioè *perché* occorresse una nuova teoria che giustifichi il linguaggio dei fisici. Ricordiamo a questo riguardo il commento sprezzante di Heaviside: “Dovrò rifiutare il mio pasto solo perché non comprendo fino in fondo i processi digestivi? No, se sono soddisfatto con il risultato”. E' chiaro che su questo tema si scontrano sensibilità molto diverse da parte di matematici, fisici, ingegneri, da parte di teorici e utilizzatori della teoria. Dobbiamo semplicemente tenerci ognuno la nostra opinione, o c'è un terreno di dialogo possibile? Nell'ultimo paragrafo (§ 5.3) proporrò qualche mia riflessione sul rapporto tra matematica e conoscenza fisica che va nella direzione di una possibile risposta.

5.1 Le trappole logiche (o psicologiche) dei vecchi racconti sulla delta

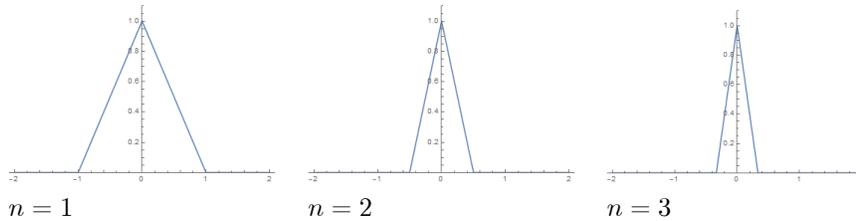
5.1.1 Impossibili scambi tra limiti e integrali

1. Dimostriamo che (come affermato tante volte in precedenza) non esiste nessuna funzione $\delta(x)$ localmente integrabile per cui valga l'uguaglianza

$$\int_{\mathbb{R}} f(x) \delta(x) dx = f(0)$$

per ogni funzione f continua (e, ad esempio, nulla fuori da un intervallo limitato). Supponiamo per assurdo che esista tale funzione $\delta(x)$, e consideriamo la successione di funzioni, continue e nulle fuori da $[-1, 1]$,

$$f_n(x) = \begin{cases} 1 - n|x| & \text{se } |x| < \frac{1}{n} \\ 0 & \text{altrimenti.} \end{cases}$$



Poiché $f_n(0) = 1$, $0 \leq f_n(x) \leq 1$, e $f_n(x) = 0$ fuori da $[-\frac{1}{n}, \frac{1}{n}]$, si avrebbe:

$$1 = f_n(0) = \int_{-1}^1 f_n(x) \delta(x) dx, \text{ quindi}$$

$$1 = \left| \int_{-1}^1 f_n(x) \delta(x) dx \right| \leq \int_{-\frac{1}{n}}^{\frac{1}{n}} |\delta(x)| dx \rightarrow 0 \text{ per } n \rightarrow \infty \quad (6)$$

perché una proprietà nota dell'integrale di Lebesgue assicura che se una funzione δ ha integrale finito su $[-1, 1]$, allora l'integrale di δ esteso a un intervallo contenuto in $[-1, 1]$ e di lunghezza sempre più piccola, tende a zero. Dunque la (6) darebbe $1 = 0$, assurdo.

2. Il *rimedio ingenuo* che si cerca è vedere la delta come un opportuno limite. In effetti ci sono molti modi di definire una successione di funzioni $\{\delta_n(x)\}_{n=1}^{\infty}$, con grafici a campana “sempre più concentrati attorno all’origine” e con le seguenti proprietà:

$$\int_{\mathbb{R}} \delta_n(x) dx = 1 \text{ per ogni } n \quad (7)$$

$$\delta_n(x) \rightarrow \begin{cases} 0 & \text{per } x \neq 0 \\ +\infty & \text{per } x = 0 \end{cases} \text{ per } n \rightarrow +\infty. \quad (8)$$

$$\int_{\mathbb{R}} f(x) \delta_n(x) dx \rightarrow f(0) \text{ per } n \rightarrow \infty \quad (9)$$

per ogni funzione f continua. Dopo di che si dice: “possiamo vedere la δ come il limite della successione δ_n ”. E’ quello che si trova scritto in tanti libri di fisica o ingegneria, è quello che faceva Cauchy nel 1815, col linguaggio degli infinitesimi anziché quello dei limiti. Ma questo risolve il problema di definire la δ ? No, perché la (8) implica che

$$\delta(x) = \begin{cases} 0 & \text{per } x \neq 0 \\ +\infty & \text{per } x = 0 \end{cases}$$

e questa funzione δ non soddisfa né

$$\int_{\mathbb{R}} \delta(x) dx = 1$$

(infatti, questa δ ha integrale nullo), né la

$$\int_{\mathbb{R}} f(x) \delta(x) dx = f(0)$$

(infatti, l'integrale vale zero per *ogni* f continua, anche per quelle che in 0 non valgono 0).

La trappola logica (o forse psicologica) di questo rimedio ingenuo è l'identificazione tra

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \delta_n(x) f(x) dx \text{ e } \int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) f(x) dx$$

e tra

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \delta_n(x) dx \text{ e } \int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) dx.$$

Ci fa comodo pensare che siccome

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \delta_n(x) f(x) dx = f(0)$$

allora sia anche

$$\int_{\mathbb{R}} \delta(x) f(x) dx = \int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) f(x) dx = f(0)$$

mentre in effetti

$$\int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) f(x) dx = 0.$$

Analogamente, *ci fa comodo pensare* che siccome

$$\int_{\mathbb{R}} \delta_n(x) dx = 1 \text{ per ogni } n,$$

e quindi anche

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \delta_n(x) dx = 1$$

allora sia anche

$$\int_{\mathbb{R}} \delta(x) dx = \int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) dx = 1$$

mentre in effetti

$$\int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) dx = 0.$$

La questione quindi è *lo scambio d'ordine, lecito o illecito, tra un integrale e un limite*.

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} \delta_n(x) f(x) dx = \int_{\mathbb{R}} \lim_{n \rightarrow \infty} \delta_n(x) f(x) dx ???$$

Esistono teoremi di analisi che affermano che sotto certe ipotesi lo scambio tra un limite e un integrale è lecito; come sempre in matematica, quando le ipotesi di un teorema del genere non sono soddisfatte, ci sono esempi in cui l'uguaglianza è effettivamente falsa. L'esempio della successione di curve a campana che tende alla delta è proprio uno di quelli in cui l'uguaglianza è falsa. Perciò giustificare

l'utilizzo di una ipotetica funzione delta che integrata contro una funzione continua restituirebbe il valore $f(0)$ mediante il limite di una successione di curve a campana significa *utilizzare lo scambio tra un limite e un integrale esattamente in un caso in cui sappiamo che questo passaggio produce un'uguaglianza falsa*.

3. Notiamo che questo è lo stesso tipo di errore che faceva Fourier nel 1822, quando scriveva

$$f(x) = \int_{-\pi}^{\pi} \delta(\alpha - x) f(\alpha) d\alpha \text{ con } \delta(x) = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right).$$

Il punto suggestivo di quest'identità è che se poniamo

$$\delta_n(x) = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^n \cos(kx) \right)$$

allora è vero che

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} \delta_n(\alpha - x) f(\alpha) d\alpha = f(x)$$

e a noi *piacerebbe poter concludere* che

$$\int_{-\pi}^{\pi} \lim_{n \rightarrow \infty} \delta_n(\alpha - x) f(\alpha) d\alpha = f(x).$$

Purtroppo però, $\lim_{n \rightarrow \infty} \delta_n(\alpha - x)$ non esiste mai, e quindi è certamente falsa l'uguaglianza

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} \delta_n(\alpha - x) f(\alpha) d\alpha = \int_{-\pi}^{\pi} \lim_{n \rightarrow \infty} \delta_n(\alpha - x) f(\alpha) d\alpha.$$

5.1.2 La necessità delle costruzioni matematiche

C'è un altro aspetto errato che voglio mettere in luce, più al fondo, in questi procedimenti di limite, e che li rende dei “tentativi disperati”. La questione è: abbiamo un oggetto misterioso che vorremmo definire, la delta, e proviamo a farlo vedendo la delta come limite di una successione di oggetti già noti, funzioni. Chiediamoci se questo procedimento ha speranza di riuscire, indipendentemente da quale sia la successione approssimante usata. Infatti, potrebbe restare il dubbio che si possa scegliere una successione di funzioni δ_n in modo più astuto di quanto proposto fin qui, e ottenere come limite proprio la delta.

La domanda cruciale è: la delta è una funzione o è un “nuovo” tipo di oggetto? (“Nuovo”, ad esempio, rispetto alla matematica che era nota nell'800). Il rigore della matematica di fine ottocento era più che sufficiente per capire che non può esistere una funzione nulla fuori dall'origine e avente integrale 1. Se dunque la delta non è una funzione, occorre *prima* costruire il nuovo sistema di oggetti più generali a cui anche la delta appartiene (ad es. le distribuzioni),

e solo poi si potrà definire la delta (con un procedimento di limite o magari un procedimento molto più semplice e diretto). Ma finché il nuovo universo di oggetti non è stato costruito, è inutile provare a definire la delta con un limite, sarebbe come volersi sollevare da terra tirandosi per i capelli. Spieghiamo questo punto con una analogia. In Analisi 1 il numero e di Nepero viene definito come il limite:

$$\lim_{n \rightarrow +\infty} \left(1 + \frac{1}{n}\right)^n .$$

Ora, i numeri $\left(1 + \frac{1}{n}\right)^n$ sono tutti numeri razionali, mentre il limite della successione, appunto il numero e , sappiamo che è irrazionale. Questo modo di definire e ha senso solo perché in precedenza, in qualche modo, si è definito l'ambiente dei numeri reali, e si sa cosa sono i numeri razionali e irrazionali. Per una persona che conoscesse solo i numeri razionali, la successione $\left(1 + \frac{1}{n}\right)^n$ semplicemente *non avrebbe alcun limite*, e la definizione di e perderebbe senso. Si può quindi usare un procedimento di limite per definire uno specifico numero irrazionale solo una volta che si sa cosa sono i numeri irrazionali. Finché si conoscono *solo* i razionali non si possono *definire* i numeri irrazionali come limiti di successioni di numeri razionali, perché finché non si conoscono gli irrazionali quelle successioni semplicemente non hanno limite.

Analogamente: non si può *definire* l'unità immaginaria come la radice quadrata di -1 , perché finché si conoscono solo i numeri reali, il numero -1 semplicemente non ha una radice; non si possono *definire* i numeri razionali come i quozienti tra interi, perché finché si conoscono solo gli interi il quoziente $\frac{5}{2}$, ad esempio, semplicemente non esiste. La “rigorizzazione dell'analisi” compiuta nell'800 ci ha insegnato la necessità di utilizzare opportune *costruzioni matematiche*, che con un certo grado di astrazione e complessità ci portano ad allargare via via l'universo di oggetti conosciuti, dagli interi ai razionali, ai reali, ai complessi; dalle funzioni alle distribuzioni. Così l'insieme dei numeri complessi si può costruire a partire dall'insieme dei numeri reali, considerando le coppie ordinate di numeri reali e definendo su queste coppie opportune operazioni di somma e prodotto (come si trova spiegato su molti testi di Analisi 1); l'insieme dei numeri reali si può costruire a partire da quello dei numeri razionali con una procedura un po' più complicata²⁷.

L'insegnamento generale di queste costruzioni è che *per dare significato a uno specifico “oggetto strano” che non rientra nell'universo che si conosce già, non ci sono scorciatoie: di solito si deve costruire un intero nuovo universo di oggetti più ampio, al cui interno poi si potrà trovare in particolare quello specifico “oggetto strano”, che a posteriori non apparirà più così strano. Cercare invece di costruire direttamente quello specifico oggetto strano vedendolo come limite di una successione di oggetti già noti apre inevitabilmente dei circoli viziosi.*

²⁷ce n'è una varietà in letteratura, dalle “sezioni” di Dedekind del 1872 agli “allineamenti decimali” del libro di testo di Analisi 1 di Pagani-Salsa del 1990, per citare solo un paio di esempi che apprezzo.

5.2 Cosa ci ha fatto guadagnare la teoria delle distribuzioni*

Vorrei dare ora qualche cenno ad alcuni settori dell'analisi matematica in cui la teoria delle distribuzioni ha avuto un forte impatto. Toccherò brevemente tre argomenti²⁸.

5.2.1 Equazioni alle derivate parziali

Come abbiamo già accennato (v. § 4.5), *ogni distribuzione è derivabile infinite volte*. Questo sembra un ottimo punto di partenza per studiare le equazioni differenziali. Ad esempio, se L è un operatore differenziale lineare a coefficienti costanti (di qualsiasi ordine), come il laplaciano, l'operatore delle onde o del calore, e u è una distribuzione, allora Lu è ben definita ed è ancora una distribuzione. Questo significa che ha senso studiare equazioni del tipo $Lu = T$ con T distribuzione assegnata, e la soluzione u che si cerca è, a priori, anch'essa una distribuzione. Ci si può chiedere se sia poi così confortante sapere che un'equazione differenziale ha una soluzione *distribuzionale*: dopo tutto, a noi piacerebbe trovare una soluzione "vera", cioè una funzione u che risolve $Lu = f$ quando f è un termine noto "vero", cioè una funzione. La teoria delle distribuzioni non ignora questo legittimo desiderio, ma spezza il problema in due sottoproblemi. Prima si dimostra che una certa equazione, sotto ipotesi molto generali, ha almeno una soluzione distribuzionale (il problema della *risolubilità*); poi si dimostra che, se il termine noto f ha una certa regolarità e l'operatore L è di un certo tipo, allora la soluzione u in effetti è più regolare di una generica distribuzione, in particolare è una funzione, tanto più regolare quanto più regolare è f (questo è il problema della *regolarità*). Così facendo, almeno in certe situazioni, alla fine si riesce a ottenere una "soluzione classica", di cui però non si sarebbe riusciti a dimostrare l'esistenza direttamente, usando solo strumenti classici.

In particolare, se L è un operatore come sopra, ha senso studiare l'equazione

$$Lu = \delta$$

(dove il termine noto è la delta di Dirac) e la soluzione di quest'equazione sarà, per definizione, la *soluzione fondamentale*, chiamiamola Γ , dell'operatore L . Il concetto di soluzione fondamentale è uno strumento centrale nello studio dei problemi sia di *risolubilità* che di *regolarità* delle soluzioni di equazioni alle derivate parziali. Anzitutto, infatti, a partire dalla conoscenza della soluzione fondamentale Γ dell'operatore L si può ottenere almeno una soluzione dell'equazione $Lu = f$ quando il termine noto f è, ad esempio, una funzione test²⁹.

²⁸Un'avvertenza: questo paragrafo risente più degli altri della limitatezza delle mie conoscenze e dei miei gusti personali. Inoltre, per forza di cose dovrò citare anche qualche concetto matematico non del tutto elementare.

²⁹La soluzione di $Lu = f$ si ottiene ponendo $u = \Gamma * f$, dove $*$ denota l'operazione di *convoluzione*. I seguenti passaggi, semplici e rigorosi all'interno della teoria delle distribuzioni, mostrano che u risolve proprio $Lu = f$:

$$Lu = L(\Gamma * f) = (L\Gamma) * f = \delta * f = f.$$

In secondo luogo, l'eventuale regolarità di Γ serve a provare risultati di regolarità per l'operatore L .

Un primo esempio storico di risultato profondo per l'analisi matematica dimostrato grazie alla teoria delle distribuzioni (ma non da Schwartz) è il teorema di Malgrange-Ehrenpreis del 1956 che afferma l'esistenza di una soluzione fondamentale per *qualsiasi* operatore differenziale lineare a coefficienti costanti.

Se ci spostiamo dagli operatori a coefficienti costanti a quelli a coefficienti variabili, le buone notizie precedenti si stemperano un po', per il seguente motivo tecnico. Sulle distribuzioni si possono fare molte operazioni, ma non tutte; ad esempio, in generale non si possono *moltiplicare tra loro* due distribuzioni. Una generica distribuzione si può moltiplicare solo per una funzione infinitamente derivabile. La conseguenza è che se L è un operatore differenziale a coefficienti variabili e u è una distribuzione, in generale possiamo garantire che Lu sia ben definita come distribuzione solo a patto che i coefficienti dell'operatore L siano infinitamente derivabili. *La teoria delle distribuzioni fornisce quindi il quadro concettuale per lo studio delle equazioni alle derivate parziali lineari a coefficienti costanti oppure variabili ma estremamente regolari.* In questo contesto si può cercare di seguire lo stesso approccio descritto sopra: dimostrare prima teoremi di esistenza (risolubilità) e poi di regolarità. In questo solco si inquadra ad esempio la ricerca svolta da **Lars Hörmander (1931 – 2012)** a partire dalla metà degli anni 1950, culminata nella medaglia Fields da lui vinta nel 1962, nella monografia [17] "Linear partial differential operators" del 1963, e quindi nella monumentale opera "The Analysis of Linear Partial Differential Operators", in 4 volumi. Il lettore che desideri qualche informazione un po' meno vaga su quest'argomento può leggere il primo capitolo (poche pagine) del testo [2], ed eventualmente consultare la bibliografia ivi citata.

Una domanda naturale che rimane aperta dalla discussione precedente è: come trattiamo gli operatori differenziali a coefficienti variabili ma non regolari? Per questi esistono altre teorie, che tipicamente fanno uso dei già citati spazi di Sobolev, ma non entreremo in questo argomento.

5.2.2 Successioni e serie di distribuzioni

Una volta che si conosce la teoria delle distribuzioni, si può definire un concetto (molto semplice) di *limite di una successione di distribuzioni*:

Definizione 12 *Data una successione di distribuzioni $\{T_n\}_{n=1}^\infty$, $T_n \in \mathcal{D}'(\mathbb{R})$ e $T \in \mathcal{D}'(\mathbb{R})$, diremo che*

$$T_n \rightarrow T \text{ in } \mathcal{D}'(\mathbb{R})$$

se

$$\langle T_n, \phi \rangle \rightarrow \langle T, \phi \rangle \text{ per } n \rightarrow \infty, \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}).$$

Questa definizione permette di recuperare il senso delle affermazioni ingenuie sulla delta come limite di successioni di curva a campana sempre più concentrate.

Ad esempio, se $\delta_n(x)$ è una successione di funzioni del tipo “nucleo di Cauchy” (3) (dove al posto del parametro α “infinitamente piccolo” poniamo $1/n$)

$$\delta_n(x) = \frac{1}{\pi} \frac{1/n}{1/n^2 + x^2} = \frac{1}{\pi} \frac{n}{1 + n^2 x^2}$$

risulta effettivamente che, per $n \rightarrow \infty$,

$$\langle T_{\delta_n}, \phi \rangle = \int_{\mathbb{R}} \frac{1}{\pi} \frac{n}{1 + n^2 x^2} \phi(x) dx \rightarrow \phi(0) = \langle \delta, \phi \rangle$$

perciò possiamo dire che la successione di distribuzioni T_{δ_n} che corrisponde a δ_n tende alla δ :

$$T_{\delta_n} \rightarrow \delta \text{ in } \mathcal{D}'(\mathbb{R})$$

o anche, parlando più alla buona,

$$\delta_n \rightarrow \delta \text{ in } \mathcal{D}'(\mathbb{R}).$$

Si noti che, come già osservato nel § 5.1, questa affermazione non serve a *definire* la δ (senza già sapere cosa sono le distribuzioni).

Dal limite di successioni di distribuzioni nasce in modo naturale la definizione di *serie di distribuzioni*. Così come la somma di una serie numerica è il limite della successione delle sue somme parziali, allo stesso modo definiremo la somma di una serie di distribuzioni:

Definizione 13 *Data una successione di distribuzioni $\{T_n\}_{n=1}^{\infty}$, $T_n \in \mathcal{D}'(\mathbb{R})$ e $T \in \mathcal{D}'(\mathbb{R})$, diremo che*

$$\sum_{n=1}^{\infty} T_n = T$$

se

$$\sum_{k=1}^n T_k \rightarrow T \text{ in } \mathcal{D}'(\mathbb{R}) \text{ per } n \rightarrow \infty,$$

cioè se

$$\left\langle \sum_{k=1}^n T_k, \phi \right\rangle \rightarrow \langle T, \phi \rangle \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}), \text{ per } n \rightarrow \infty,$$

La cosa notevole è che la definizione precedente si può riscrivere anche così:

$$\sum_{n=1}^{\infty} \langle T_n, \phi \rangle = \left\langle \sum_{n=1}^{\infty} T_n, \phi \right\rangle \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}),$$

dove cioè lo scambio tra la *serie* e l'*azione del funzionale* $\langle \cdot, \phi \rangle$ vale ora *per definizione*. E' sorprendente il confronto tra quest'identità e la cautela che è richiesta per scambiare, ad esempio, una serie e un integrale. Non solo: si dimostra (facilmente!) che

$$\left(\sum_{n=1}^{\infty} T_n = T \right) \Rightarrow \left(\sum_{n=1}^{\infty} T'_n = T' \right)$$

cioè: una serie convergente di distribuzioni si può *sempre* derivare termine a termine!

Un'applicazione sorprendente del concetto di serie di distribuzioni è la facilità con cui la teoria dà senso a certe *serie trigonometriche divergenti*. Consideriamo l'identità "folle" che scriveva Fourier nel 1822:

$$f(x) = \int_{-\pi}^{\pi} \delta(\alpha - x) f(\alpha) d\alpha$$

dove

$$\delta(x) = \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right) \text{ per } x \in [-\pi, \pi].$$

Per semplicità, consideriamo le uguaglianze precedenti per $x = 0$ e $f = \phi \in \mathcal{D}(-\pi, \pi)$ (cioè la ϕ è una funzione test nulla fuori dall'intervallo $(-\pi, \pi)$). L'affermazione di Cauchy quindi diventa:

$$\phi(0) = \int_{\mathbb{R}} \delta(x) \phi(x) dx$$

con $\delta(x)$ come sopra. Riscriviamo quest'identità col linguaggio delle distribuzioni:

$$\left\langle \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right), \phi \right\rangle = \phi(0) = \langle \delta, \phi \rangle$$

(dove ora il simbolo δ indica la distribuzione delta di Dirac, e non la "funzione" $\delta(x)$ scritta poco sopra), cioè

$$\frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right) = \delta. \quad (10)$$

Si sta ora affermando che una certa serie (di funzioni localmente integrabili, che perciò sono anche distribuzioni) converge e ha per somma la distribuzione δ di Dirac. E' vero questo? Per definizione, questo significa:

$$\left\langle \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^n \cos(kx) \right), \phi \right\rangle \rightarrow \phi(0) \text{ per } n \rightarrow \infty \text{ per ogni } \phi \in \mathcal{D}(-\pi, \pi),$$

ossia

$$\begin{aligned} & \int_{-\pi}^{\pi} \frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^n \cos(kx) \right) \phi(x) dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(x) dx + \sum_{k=1}^n \int_{-\pi}^{\pi} \phi(x) \cos(kx) dx \rightarrow \phi(0) \text{ per } n \rightarrow \infty \text{ per ogni } \phi \in \mathcal{D}(-\pi, \pi), \end{aligned}$$

e questo in effetti può essere dimostrato. Quindi la (10) è vera. Riflettiamo sul percorso logico:

1. La serie di funzioni

$$\frac{1}{\pi} \left(\frac{1}{2} + \sum_{k=1}^{\infty} \cos(kx) \right) \quad (11)$$

diverge in ogni punto x .

2. Però, per ogni $\phi \in \mathcal{D}(-\pi, \pi)$ fissata, la serie numerica

$$\sum_{k=1}^{\infty} \int_{-\pi}^{\pi} \phi(x) \cos(kx) dx$$

in effetti converge.

3. Questo permette di dare senso alla serie (11): come serie di funzioni è divergente, ma come serie di distribuzioni converge.

In modo analogo si può dimostrare che *convergono nel senso delle distribuzioni* tutte le serie trigonometriche del tipo:

$$\sum_{k=1}^{\infty} \{a_k \cos(kx) + b_k \sin(kx)\}$$

quando $\{a_k\}, \{b_k\}$ sono successioni di coefficienti numerici “lentamente crescenti”, ossia che tendono a infinito come una potenza k^α per qualche $\alpha > 0$. Questo fatto è *soprendente*, perché viste come *serie di funzioni* queste sono un caso disperato; ed è anche *molto utile*, perché queste serie sono esattamente quelle che si ottengono quando si parte dalla serie di Fourier (convergente) di una funzione “ragionevole” e la si *deriva termine a termine* un certo numero di volte.

Esempio 14 *Nello studio elementare delle serie di Fourier si dimostra che:*

$$x^2 = \frac{\pi^2}{3} + 4 \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \cos(kx) \text{ per ogni } x \in [-\pi, \pi],$$

dove la serie a secondo membro converge (in tutti i sensi ragionevoli). Proviamo a derivare “ingenuamente” ambo i membri di quest’identità, una volta

$$2x = 4 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sin(kx)$$

e poi un’altra ancora

$$2 = 4 \sum_{k=1}^{\infty} (-1)^{k+1} \cos(kx).$$

Abbiamo ottenuto un’assurdità: la serie a secondo membro non converge in nessun punto; come può la sua somma essere uguale a 2 per ogni x ? Se poi deriviamo ancora una volta:

$$0 = 4 \sum_{k=1}^{\infty} (-1)^k k \sin(kx)$$

l'uguaglianza si fa ancora più assurda. Eppure, nel senso delle distribuzioni, la serie converge e le uguaglianze scritte sono tutte vere, in $\mathcal{D}'(-\pi, \pi)$. Esplicitamente questo significa:

$$\sum_{k=1}^{\infty} (-1)^k k \int_{-\pi}^{\pi} \phi(x) \sin(kx) dx = 0 \text{ per ogni } \phi \in \mathcal{D}(-\pi, \pi).$$

La teoria delle distribuzioni fornisce quindi un quadro solido entro cui trattare identità che coinvolgono le serie trigonometriche, derivarle termine a termine un numero qualsiasi di volte, trovando uguaglianze (tra distribuzioni) sensate e vere, laddove leggendo uguaglianze tra funzioni si troverebbero solo assurdità. In certi casi significativi, queste uguaglianze “ardite” costituiscono i passaggi intermedi di una dimostrazione in cui il risultato finale, invece, ha un senso anche in termini di uguaglianze di funzioni. Sono questi i modi in cui la teoria si dimostra uno strumento potente per ottenere rigorosamente risultati che, nel quadro delle funzioni, non sarebbero altrettanto facili da ottenere.

5.2.3 Trasformata di Fourier

La teoria delle distribuzioni permette di generalizzare il concetto di *trasformata di Fourier* di una funzione. Se $f : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione integrabile, si definisce la sua trasformata di Fourier

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x \xi} dx.$$

Se la variabile x ha il significato di *tempo*, f si può vedere come un *segnale*, che la trasformata di Fourier va ad analizzare nelle sue infinite *frequenze* ξ .

Purtroppo, nessuna funzione elementare (potenze, esponenziali, funzioni trigonometriche...) è integrabile su tutta la retta; lo stesso succede per le funzioni periodiche (non identicamente nulle). Ci sono quindi molti “segnali” interessanti per i quali la definizione standard di trasformata di Fourier perde significato. La teoria delle distribuzioni consente di rimediare.

Se $T \in \mathcal{D}'(\mathbb{R})$, si vuole definire la sua trasformata di Fourier \widehat{T} . L'idea, per indovinare la definizione ragionevole di un'operazione nell'ambito delle distribuzioni, è sempre la stessa: nel caso particolare in cui $T = T_f$ con f funzione integrabile, per la quale sappiamo già cos'è la trasformata \widehat{f} , vorremmo ritrovare

$$\widehat{(T_f)} = T_{\widehat{f}}.$$

In termini di funzioni test, questo significa richiedere che valga:

$$\langle \widehat{(T_f)}, \phi \rangle = \langle T_{\widehat{f}}, \phi \rangle$$

ma

$$\langle T_{\widehat{f}}, \phi \rangle = \int_{\mathbb{R}} \widehat{f}(y) \phi(y) dy$$

e per una proprietà nota (e di facile dimostrazione) se f e ϕ sono due funzioni integrabili, l'ultimo integrale scritto è uguale a

$$\int_{\mathbb{R}} f(y) \widehat{\phi}(y) dy,$$

che si può vedere come

$$\langle T_f, \widehat{\phi} \rangle.$$

In definitiva, noi vogliamo che la definizione di trasformata di Fourier di una distribuzione, quando la si applichi alle distribuzioni di tipo T_f che “provengono” da funzioni integrabili, soddisfi l'identità

$$\langle \widehat{(T_f)}, \phi \rangle = \langle T_f, \widehat{\phi} \rangle.$$

Allora è naturale definire la trasformata di una distribuzione T qualsiasi al modo seguente:

$$\langle \widehat{T}, \phi \rangle = \langle T, \widehat{\phi} \rangle \text{ per ogni } \phi \in \mathcal{D}(\mathbb{R}).$$

Questa è l'unica definizione ragionevole da dare. Purtroppo però a questo punto nasce un problema tecnico: se $\phi \in \mathcal{D}(\mathbb{R})$, la trasformata $\widehat{\phi}$ è sempre una funzione infinitamente derivabile, ma non è *mai* nulla fuori da un intervallo limitato. In termini fisici: non esistono segnali che siano contemporaneamente *di durata finita e a banda limitata*. Questo significa che $\widehat{\phi}$ non è una funzione test, perciò il numero $\langle T, \widehat{\phi} \rangle$ potrebbe non essere definito.

Questo problema è reale, ma superabile senza gravi danni. Si tratta di rassegnarsi al fatto che non tutte le operazioni, nell'ambito delle distribuzioni, funzionano “senza alcuna restrizione” come accade ad esempio per la derivata. Nel caso della trasformata di Fourier, esistono distribuzioni trasformabili e altre non trasformabili. Si definisce una certa classe di distribuzioni, dette *distribuzioni temperate*, per le quali la definizione precedente funziona bene. Per definire le distribuzioni temperate, si tratta di allargare lo spazio delle funzioni test, passando dallo spazio $\mathcal{D}(\mathbb{R})$ allo “spazio di Schwartz”³⁰ $\mathcal{S}(\mathbb{R})$. Le distribuzioni temperate sono allora, per definizione, i funzionali lineari continui (in un senso da precisare) su questo spazio $\mathcal{S}(\mathbb{R})$. Sorvolando sui dettagli tecnici, facciamo solo un esempio:

Esempio 15 *La delta di Dirac possiede trasformata di Fourier, e si ha:*

$$\widehat{\delta} = 1.$$

Infatti:

$$\langle \widehat{\delta}, \phi \rangle = \langle \delta, \widehat{\phi} \rangle = \widehat{\phi}(0) = \int_{\mathbb{R}} \phi(y) dy = \langle 1, \phi \rangle,$$

³⁰Una funzione appartiene allo spazio $\mathcal{S}(\mathbb{R})$ se è infinitamente derivabile, non è necessariamente nulla fuori da un intervallo limitato, ma tende a zero all'infinito, insieme alle sue derivate di ogni ordine, più rapidamente di qualunque potenza a esponente negativo. Il bello di questo spazio è che se $\phi \in \mathcal{S}(\mathbb{R})$, allora anche $\widehat{\phi} \in \mathcal{S}(\mathbb{R})$. Naturalmente ci sarebbero molti altri dettagli da precisare, ma qui non lo faremo.

perciò $\widehat{\delta} = 1$. Quest'esempio si può interpretare intuitivamente a questo modo: se un segnale è totalmente concentrato in un istante, sarà distribuito in modo perfettamente uniforme in tutte le frequenze.

Con un ragionamento simile si dimostra che, reciprocamente,

$$\widehat{1} = \delta,$$

esempio significativo perché in questo caso $f(x) = 1$ è una funzione, che però non è integrabile. La sua trasformata di Fourier esiste in senso distribuzionale, ma non è una funzione.

Si può facilmente immaginare come il fatto di aver “liberato” la teoria della trasformata di Fourier dalla pesante richiesta di integrabilità globale delle funzioni abbia reso la teoria molto più flessibile.

5.3 Matematica e conoscenza fisica

Veniamo quindi all'ultima domanda lasciata aperta, cioè se e perché occorresse proprio “una teoria matematica nuova che giustifichi, in forma modificata, il linguaggio dei fisici”, come scriveva Schwartz. Detto in altre parole, fino a che punto è giusto insistere sul rigore matematico anche quando la matematica viene usata “in casa dei fisici”, o reciprocamente, fino a che punto sono accettabili, in campo applicato, passaggi matematici motivati su una base euristica ma privi di rigore?

Proverò a rispondere allargando il discorso al tema del rapporto tra matematica e conoscenza fisica.

Gli utilizzatori della matematica si riferiscono ad essa molto spesso con il termine *strumento*: “la matematica è uno strumento”, dicono. Spero possiate immaginare che a un matematico questo modo di esprimersi non piaccia. Per chi dedica la propria vita professionale alla matematica, qualificarla come strumento è qualcosa che sta molto stretto. Io preferisco dire che la matematica è un *metodo di conoscenza*. Se proprio volete usare la parola strumento, dite almeno *strumento di conoscenza*.

Ma cosa qualifica il metodo matematico, almeno da Euclide in poi? La *dimostrazione*, che ricava nuove verità da assunti già accettati. La dimostrazione è il cuore della specificità di quel metodo di conoscenza che chiamiamo matematica. Un procedimento che con un numero finito di passaggi logici è in grado di convincere noi stessi e altre persone della validità di infiniti asserti, che una vita intera non basterebbe a “controllare empiricamente”. Un procedimento che, di certe cose che sono vere, ci fa capire *perché* sono vere. Mi piace citare a questo proposito il prossimo passo, volutamente enfatico ma significativo, dal saggio “L'esperienza matematica” di Davis e Hersh:

“La dimostrazione serve a molti propositi simultaneamente. Essendo esposta allo scrutinio e al giudizio di un pubblico sempre nuovo, la dimostrazione è soggetta a un costante processo di critica e rivalidazione. Errori, ambiguità e fraintendimenti sono ripuliti nel

corso della costante esposizione. La dimostrazione è rispettabilità. La dimostrazione è il sigillo dell'autorità. La dimostrazione, nei suoi migliori esempi, accresce la comprensione rivelando il cuore del problema. La dimostrazione suggerisce nuova matematica. La dimostrazione è la potenza matematica, la scarica elettrica di questa materia che rivitalizza le statiche asserzioni dei teoremi. Infine, la dimostrazione è rituale, celebrazione del potere della ragione pura. Un tale esercizio di rassicurazione può essere estremamente necessario in vista di tutti i problemi in cui ci getta il pensiero logico.” [5, p.151]

Ma questo aspetto metodologico, così centrale per la matematica in se stessa, è anche il contributo, o forse dovremmo dire il regalo, che la matematica dà alla fisica, come già diceva Galilei: lo studio della natura si fonda, diceva, su “*sensate esperienze*” e “*dimostrazioni matematiche*”.

“(…) così si costuma e conviene nelle scienze le quali alle conclusioni naturali applicano le dimostrazioni matematiche, come si vede ne i prospettivi, negli astronomi, ne i meccanici, ne i musici ed altri, li quali con sensate esperienze confermano i principii loro, che sono i fondamenti di tutta la seguente struttura (...)”³¹

La matematica è una delle due gambe su cui cammina il sapere scientifico. Ma se vogliamo che questa gamba non sia zoppa, la dimostrazione *dev'essere una dimostrazione*.

La tautologia precedente apre in realtà una questione, e cioè chi o cosa decida se una presunta dimostrazione sia davvero tale. Certamente i *canoni del rigore* si sono evoluti nella storia della matematica, sotto la spinta di vari fattori, tra cui le proprietà degli stessi oggetti matematici che si indagavano. Così, ad esempio, Newton nel '700 poteva permettersi di giocare con le *serie di potenze* senza troppi riguardi di rigore e cascava sempre in piedi, perché in effetti una serie di potenze, se converge in un intervallo, si può anche derivare e integrare termine a termine tutte le volte che si vuole. I matematici dell'800 non potevano permettersi di giocare con la stessa spregiudicatezza con le serie trigonometriche (di Fourier), perché appena si prova a farlo spuntano ovunque affermazioni contraddittorie: serie che non convergono in nessun punto, ad esempio. Cercando di capirci qualcosa di più, i matematici dell'800 hanno dovuto dare una definizione di funzione, una definizione di limite, hanno dovuto rendere rigoroso il calcolo infinitesimale, hanno creato la teoria degli insiemi, e così via. I canoni del rigore si evolvono quindi anche a motivo del tipo di oggetti in cui ci imbattiamo, che segnalano i limiti di validità di affermazioni che precedentemente davamo per assolutamente vere, un po' come è successo in fisica esplorando l'universo sulle scale molto grandi e molto piccole. Tuttavia, una volta che il problema da cui

³¹Galileo Galilei: Discorsi e dimostrazioni matematiche intorno a due nuove scienze attenenti alla meccanica e i movimenti locali, 1638. Giornata Terza, sezione “Del moto naturalmente accelerato”, Salviati. (Ma concetti simili sono certamente espressi da Galilei anche in altri passi).

nasce l'esigenza di rigore si è posto, ignorarlo significa rinunciare alla conoscenza che la matematica può darci.

Se vogliamo che la matematica sia uno strumento di conoscenza quando la applichiamo alla fisica, dobbiamo poterci fidare *a priori* delle sue deduzioni: dobbiamo poter credere che, partendo da premesse fisiche assodate, se un ragionamento matematico rigoroso ha una certa conseguenza fisica, questa debba essere vera. L'uso disinvolto della matematica in certe applicazioni normalmente si giustifica perché "tanto si sa già cosa è ragionevole aspettarsi". Ma così facendo ci si perde il meglio, cioè:

1. la fiducia *a priori* nelle sue conclusioni;
2. la possibilità di essere talvolta anche stupiti da una conseguenza inaspettata, ma ostinatamente vera.

Se vogliamo che la matematica continui a darci nuova conoscenza, dobbiamo usarla rispettando il suo metodo, il metodo dimostrativo, e non addomesticarla per "far tornare" quello che già ci aspettiamo.

Qualche lettura per approfondire

La seguente bibliografia raccoglie tutti i testi citati nell'articolo o da me consultati. Diversi dei testi più antichi sono consultabili online. Qui mi limito a suggerire qualche lettura per chi vuole approfondire l'aspetto storico della delta e della teoria delle distribuzioni, oppure per chi cerca dei testi su cui studiare la teoria stessa.

A. *Sulla preistoria e nascita della teoria delle distribuzioni* segnalo:

Il testo di Lützen [18]; il breve saggio di Garding *The impact of distributions in analysis*, contenuto in [10]; l'introduzione del libro di Schwartz [24].

B. *Per studiare la teoria delle distribuzioni, segnalo tra i tanti alcuni titoli, suddivisi in tre tipologie di libro, in ordine dai più impegnativi (e approfonditi) ai meno impegnativi:*

a. Testi dedicati interamente (o ampiamente) alla teoria delle distribuzioni, che fanno uso di concetti di topologia e hanno un taglio matematico teorico: oltre al testo originale di Schwartz [24], segnalo Trèves [29], Rudin [20, Part II].

b. Testi dedicati interamente (o ampiamente) alla teoria delle distribuzioni, che *non fanno uso* di concetti astratti di topologia e hanno un taglio più rivolto all'ingegneria: Duistermaat, Kolk [8], Teodorescu, Kec, Toma [28], Strichartz [27].

c. Qualche libro di testo di matematica per l'ingegneria che dedica uno-due capitoli alla teoria delle distribuzioni: segnalo, tra i tanti, Gasquet, Witomski [11, Chaps. VIII-IX], Salsa [21, Cap.7, §§ 7.1-7.4.], e il mio [1, Cap.9].

Riferimenti bibliografici

- [1] M. Bramanti: Metodi di Analisi Matematica per l'Ingegneria. Esculapio, 2019.
- [2] M. Bramanti: An invitation to hypoelliptic operators and Hörmander's vector fields. SpringerBriefs in Mathematics. Springer, Cham, 2014.
- [3] N. Bourbaki: Éléments de mathématique. Part I. Les structures fondamentales de l'analyse. Livre I. Théorie des ensembles (Fascicule de résultats). Actualités Scientifiques et Industrielles, No. 846. Hermann & Cie., Paris, 1939. 51 pp.
- [4] A. L. Cauchy: Théorie de la propagation des ondes à la surface d'un fluide pesant, d'une profondeur indéfinie. Mém. Acad. Roy. Sci. Inst. France, Sci. Math. et Phys, 1 (1827) = Oeuvres Complètes, Paris, 1882-1958, 1 (1), pp.5.-318. *Consultabile online a:* http://sites.mathdoc.fr/cgi-bin/oeitem?id=OE_CAUCHY_1_1_5_0
- [5] P. J. Davis, R. Hersh: The mathematical experience. Mariner Books, Houghton Mifflin Company, Boston, New York, 1998.
- [6] P. A. M. Dirac: The principles of quantum mechanics. First edition: 1930. Fourth edition 1958, Oxford University Press, reprinted in 1967.
- [7] G. Doetsch: Introduction to the theory and application of the Laplace transformation. Translated from the second German edition by Walter Nader. Springer-Verlag, New York-Heidelberg, 1974. vii+326 pp. *Opera originale:* Theorie und Anwendung der Laplace-Transformation. By G. Doetsch. (Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen, vol. 47.) Berlin, Springer, 1937. 8+436 pp.
- [8] J. J. Duistermaat, J. A. C. Kolk: Distributions. Theory and applications. (Translated from the Dutch by J. P. van Braam Houckgeest). Cornerstones. Birkhäuser Boston, Inc., Boston, MA, 2010.
- [9] J. B. J. Fourier, The analytical theory of heat. Cambridge Library Collection. Cambridge University Press, 2009. L'opera originale di Fourier, *Théorie Analytique de la Chaleur*, è del 1822.
- [10] L. Gårding: Some points of analysis and their history. University Lecture Series, 11. American Mathematical Society, Providence, RI; Higher Education Press, Beijing, 1997. viii+88 pp.
- [11] C. Gasquet, P. Witomski: Fourier analysis and applications. Filtering, numerical computation, wavelets. Texts in Applied Mathematics, 30. Springer-Verlag, New York, 1999.
- [12] G. Green: An Essay on the Application of mathematical Analysis to the theories of Electricity and Magnetism, 1828. *Consultabile online a:* <https://arxiv.org/pdf/0807.0088.pdf>

- [13] M. G. Katz, D. Tall: A Cauchy-Dirac delta function. *Found. Sci.* 18 (2013), no. 1, 107–123.
- [14] O. Heaviside: On operators of physical mathematics. Part I. *Proc. Roy. Soc. London*, 52 (1893), pp.504-529. *Consultabile online a:* <https://royalsocietypublishing.org/doi/pdf/10.1098/rspl.1892.0093>
- [15] O. Heaviside: *Electromagnetic Theory. vol.I.* Chelsea Publishing Company. 1971. (Opera originale del 1893).
- [16] O. Heaviside: *Electromagnetic Theory. vol.II.* Chelsea Publishing Company. 1971. (Opera originale del 1899). *Consultabile online a:* <https://archive.org/details/electromagnetict02heavrich>
- [17] L. Hörmander: *Linear partial differential operators. Die Grundlehren der mathematischen Wissenschaften, Bd. 116* Academic Press, Inc., Publishers, New York; Springer-Verlag, Berlin-Göttingen-Heidelberg 1963.
- [18] J. Lützen: The prehistory of the theory of distributions. *Studies in the History of Mathematics and Physical Sciences*, 7. Springer-Verlag, New York-Berlin, 1982 viii+232 pp.
- [19] S. S. Petrova: Heaviside and the development of the symbolic calculus. *Arch. Hist. Exact Sci.* 37 (1987), no. 1, 1–23.
- [20] W. Rudin: *Functional analysis. Second edition. International Series in Pure and Applied Mathematics.* McGraw-Hill, Inc., New York, 1991. xviii+424 pp.
- [21] S. Salsa: *Equazioni a derivate parziali. Metodi, modelli e applicazioni.* 3a ed. Springer, 2016.
- [22] L. Schwartz: *Théorie des distributions. Tome I. Actualités Sci. Ind., no. 1091 = Publ. Inst. Math. Univ. Strasbourg 9.* Hermann & Cie., Paris, 1950.
- [23] L. Schwartz: *Théorie des distributions. Tome II. Actualités Sci. Ind., no. 1122 = Publ. Inst. Math. Univ. Strasbourg 10.* Hermann & Cie., Paris, 1951.
- [24] L. Schwartz: *Théorie des distributions. Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X. Nouvelle édition, entièrement corrigée, refondue et augmentée.* Hermann, Paris 1966 xiii+420 pp.; reprinted in 1978.
- [25] L. Schwartz: Généralisation de la notion de fonction, de dérivation, de transformation de Fourier et applications mathématiques et physiques. (French) *Ann. Univ. Grenoble. Sect. Sci. Math. Phys. (N.S.)* 21, (1945), 57–74 (1946).

- [26] S. L. Sobolev: Méthode nouvelle à résoudre le problème de Cauchy pour les équations linéaires hyperboliques normales. *Mat. Sb.*, 1 (43) (1936), pp.39-71.
- [27] R. S. Strichartz: A guide to distribution theory and Fourier transforms. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.
- [28] P. P. Teodorescu, W. W. Kecs, A. Toma: Distribution theory. With applications in engineering and physics. Wiley-VCH Verlag, Weinheim, 2013.
- [29] F. Trèves: Topological vector spaces, distributions and kernels. Unabridged republication of the 1967 original. Dover Publications, Inc., Mineola, NY, 2006.
- [30] J. von Neumann: Mathematical Foundations of Quantum Mechanics. Princeton University Press, 1955. (Traduzione inglese dell'opera originale: *Matematische Grundlagen der Quantenmechanik*, 1932).

Marco Bramanti
Dipartimento di Matematica, Politecnico di Milano
Via Bonardi 9. 20133 Milano
marco.bramanti@polimi.it
<http://www1.mate.polimi.it/~bramanti/>