LABORATORI NAZIONALI DI FRASCATI
SIS – Pubblicazioni

LNF–07/ 28 (IR)
3 Dicembre 2007

Avvertenza

Con le piccole modifiche del caso, i due articoli (I,II) che seguono sono l'"Introduzione" e l'ultima "Appendice Generale" con le quali si apre, e rispettivamente si chiuderà, un libro sui "Fondamenti della Fisica Matematica Macroscopica" che l'autore sta attualmente scrivendo. Tra di esse, al momento sono state redatte circa settecento pagine. Ovviamente (II) va letto dopo (I).

Notice

With slight modifications, the two following articles (I,II) are the "Introduction" and the last "General Appendix" which opens, and respectively will close, a book on the "Foundations of Macroscopic Mathematical Physics" presently under progress by the same author. In between, some seven hundred pages have been written to date. Of course, (II) should be read after (I).

I – ALCUNE RIFLESSIONI GENERALI SULLE TEORIE FISICO-MATEMATICHE [1]

Camillo Lo Surdo

Laboratori Nazionali di Frascati dell'INFN, Frascati

I.1) RUDIMENTI DI TEORIA DEI MODELLI EMPIRICI

L'elaborazione di una **teoria fisico-matematica** (in questo articolo (I), "tfm") è un processo molto complicato, difficilmente descrivibile in termini generali e in un linguaggio non tecnico. Per un verso solitaria e imperscrutabile in certi suoi passaggi decisivi, e per l'altro legata alla collaborazione di individui con competenze e ruoli diversi (quindi in qualche misura inevitabilmente "pubblica"), una tale impresa articolata e improgrammabile lascia nondimeno distinguere al suo interno momenti, o tipi di attività, che diremo qui **di modellazione** e rispettivamente **di matematizzazione**. Nei primi, l'attenzione è rivolta alla specifica realtà fisica con l'intento di riprodurla in una sua rappresentazione idealizzata e in qualche modo organizzata, o come si dice comunemente, *salvo che da parte dei logici*, [2] appunto in un suo **modello**; mentre nei secondi si tende a sviluppare tale modello in senso matematico, attribuendogli così una capacità significante rigorosa, e in particolare un'attitudine a "prevedere" nel senso dell'implicazione "**se … allora**".

Alquanto approssimativamente, l'attività di modellazione (risp. di matematizzazione) è quella che nel linguaggio filosofico è chiamata **osservativa/induttiva** (risp. **deduttiva**). Il **ragionamento induttivo** è tentativamente descritto nella prossima sezione, mentre il

[1] Il testo di questo primo articolo (I) non è rigidamente consequenziale: le definizioni in senso stretto sono quasi del tutto assenti, e buona parte della terminologia tecnica è usata con una certa libertà, presupponendo nel lettore una sufficiente dimestichezza con essa ed i relativi concetti. In entrambi gli articoli (I,II) il carattere **neretto** serve per lo più ad evidenziare nozioni/espressioni di particolare interesse contestuale al loro primo apparire; oppure, quando è il caso, alcune vere e proprie definizioni. Il carattere *corsivo* ha invece funzione di enfatizzatore. Le citazioni letterali sono sempre riportate tra «   ». Le due Bibliografie Generali sono limitate a pochi riferimenti di base, dai quali si potrà partire per eventualmente estenderle.

[2] Con una giustificabile inversione di ruoli, e accentuandone al massimo la valenza formale, i logici chiamano infatti "modello" una realtà di cose, anche soltanto pensate, che sia immagine dell'oggetto principale della loro attenzione, cioè del sistema o teoria formale, fornendone una possibile "interpretazione". In altre parole, i modelli dei logici − che tecnicamente si dicono "semantici" − sono immagini dei loro segni in un mondo di *cose* (anche astratte); mentre i modelli dei fisici − che si dicono "matematici" o "teorici" − sono immagini di ambiti della realtà fenomenica in un mondo di *simboli* matematici. Qui abbiamo usato i termini "segno" e "simbolo" dando al secondo, a differenza del primo e alla luce dell'etimo, una qualche carica di concretezza significante. Si noti anche come nei due casi (modelli semantici vs. modelli teorici) si abbiano "flussi interpretativi" *di verso opposto*.

ragionamento deduttivo è quello standard delle teorie matematiche, o dei sistemi formali che auspicabilmente le fondano. Secondo una descrizione più accurata, l'attività di modellazione può includere (e tipicamente include) l'uso di deduzioni − quasi sempre a livello inconscio, e nella interpretazione intuitiva −, e precisamente di deduzioni interne a certe "prototeorie" fondamentali comuni a gran parte delle tfm. Sarà comodo dire **transinduttivi** i ragionamenti che, includendo tali "protodeduzioni", concorrono all'attività di modellazione, che può dunque considerarsi osservativa e transinduttiva. In presenza di una sufficiente quantità di dati empirici riproducibili ed autoconsistenti, una tfm è il possibile risultato di una complessa interazione tra i sopraddetti tipi di attività di modellazione e di matematizzazione: un duro ma esaltante esercizio a più mani fatto di osservazioni, ipotesi, esperimenti più o meno mirati o cruciali, misure, analisi di risultati, stime di errori, abbozzi e sondaggi teoretici, invenzioni e sviluppi matematici ad hoc, verifiche, modifiche, catastrofi, ripudi, ricominciamenti, … e molto altro. [3]

Ma cosa distingue definitivamente una tfm (o nel linguaggio corrente, una **scienza esatta**, o **scienza in senso stretto**, la greca επιστη'μη) da quelle che comunemente chiamiamo "scienze" tout court? E in quale preciso senso una scienza esatta è più "solida", come di norma si sostiene, di una scienza generica? Una discussione approfondita di queste questioni allargherebbe inaccettabilmente il discorso, allontanandoci troppo dal nostro tema principale. Limitandoci quindi ad illustrarne gli aspetti ad esso più rilevanti, cominceremo col ricordare che il concetto stesso di scienza − con il suo approssimativo significato di "attività di 'comprensione' (understanding) di una classe di fenomeni naturali, ispirata alla *universalizzazione contestuale e alla sintesi*, e finalizzata ad una '*capacità predittiva*' relativa a quella classe di fenomeni" [4] − è piuttosto recente, essendo stato introdotto soltanto nel XVIII secolo circa; prima di allora, esso si immergeva in quello di filosofia, e in particolare in quello di "filosofia della natura" ("philosophia naturalis").

Ovviamente le definizioni sopra riportate, e le analoghe immaginabili, sono troppo concise per non risultare vaghe, o comunque più vaghe di quanto desiderabile; ma ogni tentativo inteso a

---

[3] Non si afferma qui che la rappresentazione teorica dell'oggetto fisico considerato sia necessariamente unica; ma a fronte di rappresentazioni *predittivamente equivalenti* dello stesso oggetto, di fatto la scelta privilegia sempre la *più semplice*: «L'esperienza ci lascia una libera scelta, ma la guida aiutandoci a discernere il cammino più comodo» (H. Poincaré, "La Science et l'Hypothèse", 1903). Una ben nota immagine precorritrice di questa norma "economica" è quella del cosiddetto "rasoio di Occam". Beninteso, attribuita ad una rappresentazione teorica la qualifica di "semplice" è in parte opinabile, e ciò può dar luogo ad una certa dose di "convenzionalismo" delle scienze esatte, ovvero all'idea che alcuni dei principi che sono alla base della associata tfm abbiano la natura di convenzioni. Tra i grandi fisici-matematici di ogni tempo, Poincaré fu uno dei più inclini a questa visione convenzionalistica (entro ovvi e ristretti limiti) della sua disciplina.

[4] Sarà utile riportare un'autorevole amplificazione di questa descrizione o definizione informale, quella che si trova alla voce "science" della Micropædia della Encyclopædia Britannica (15ma ed., rist. 1997), di moderna ispirazione positivista: «Any of various intellectual activities concerned with the physical world and its phenomena, and entailing unbiased observations and systematic experimentation. In general, a science involves a pursuit of knowledge covering *general* truths or the operations of *fundamental* laws.» (corsivi nostri). Come ci aspettiamo e meglio comprenderemo, una tfm soddisfa la descrizione quotata. Beninteso la capacità "pre-dittiva" nominata nel testo non è necessariamente correlata ad un ordinamento temporale.

precisarle sostanzialmente apre scenari di scoraggiante molteplicità. Sembra insomma necessario riconoscere – questa è almeno la nostra opinione – che la questione di cosa si debba *esattamente* intendere per quella che oggi chiamiamo "scienza" senza altri attributi non ammetta una risposta al contempo abbastanza *semplice* e abbastanza *rigorosa*. Inoltre è legittimo chiedersi quale possa essere il fine pratico di una tale ricerca, soprattutto tenendo conto che essa e i suoi esiti interessano assai poco alla grande maggioranza dei moderni scienziati; più o meno nello stesso senso in cui da sempre, ma certo più legittimamente, poco interessa o interesserebbe ai musicisti una discussione filosofica sulla natura specifica della musica. [5]

La situazione è diversa, e in linea di principio altrimenti dominabile, quando ci si limiti alla considerazione delle scienze esatte, ovviamente partendo da quelle esistenti e abbastanza consolidate: ad esempio la geometria euclidea, la meccanica newtoniana, l'elettromagnetismo classico, la termodinamica razionale (diciamo, come presentata da Truesdell [6]), la teoria della relatività, la meccanica quantistica, l'elettrodinamica quantistica, … e via elencando. [7] Chi di esse affronta seriamente lo studio, applicandosi soprattutto alla soluzione di problemi specifici e concreti (in parole povere, "facendo esercizi"), presto o tardi ne percepisce quel peculiare carattere comune che poi finirà con l'assimilare come sua vera e propria seconda natura di addetto ai lavori. Il problema di collocare quelle teorie in un comune quadro definitorio trova infatti una risposta precisa; ma purtroppo non è semplice spiegare in cosa consista esattamente questa risposta, nemmeno ad un lettore di cultura fisico-matematica relativamente evoluta.

Dedicheremo gran parte di questa Sez. I.1 a tal fine, anche se ciò potrà sembrare un ingresso un po' brusco nel "cuore filosofico" della materia. Il punto è che i corsi di laurea in Fisica non prevedono la Logica Matematica nei loro piani di studio; mentre una buona familiarità del lettore con questa disciplina, sia nei suoi aspetti sintattici che semantici, faciliterebbe di molto il nostro obiettivo didattico. In particolare sarebbe auspicabile una certa conoscenza di quell'ormai vasto settore della logica moderna che va sotto il nome di **Teoria Classica dei Modelli** ("classica" in quanto associata alla logica classica); e quindi della altrettanto classica **Teoria degli Insiemi**, sia intuitiva che assiomatica. [8] Se così non fosse, il lettore potrà acquisire le nozioni e i concetti propedeutici all'intelligenza di questa sezione I.1 consultando qualche testo di base sull'argomento (vedi Bibliografia Generale).

---

[5] Di una possibile definizione dell'attività "scientifica" si occupa, come è naturale, la Filosofia della Scienza. A proposito della supponenza degli scienziati verso questo tipo di problemi, viene alla mente il sarcastico aforisma di R Feynman, secondo il quale «la filosofia della scienza serve agli scienziati più o meno quanto l'ornitologia serve agli uccelli».

[6] C. Truesdell, "Rational Thermodynamics", McGraw-Hill (1969).

[7] Che tali tfm non siano in effetti così separate come può apparire a prima vista è a sua volta un tema centrale di indagine della fisica matematica.

Inizieremo dunque con una introduzione "a grandi linee" alla teoria dei modelli. Esprimendoci informalmente, questa teoria mira a creare un ponte da una data "struttura insiemistica" astratta Σ verso un conveniente sistema formale (sf) del 1° ordine S, nella forma di una *iniezione* J che trasformi gli oggetti di Σ in corrispondenti oggetti formali omologhi (≡ della stessa natura): ad esempio, gli "enunciati" di Σ in corrispondenti enunciati formali di S, e in particolare, gli enunciati cosiddetti "atomici" (≡ logicamente non decomponibili in enunciati più semplici) di Σ in corrispondenti enunciati atomici formali di S. Sarà comodo dire che J è una iniezione "di [Σ] in [S]", intendendo che [Σ] rappresenti il prodotto cartesiano di classi di oggetti omologhi di Σ, e [S] il prodotto cartesiano delle classi degli oggetti omologhi corrispondenti di S; e scrivere dunque, con l'usuale notazione,

(1) $J: [\Sigma] \to_{ini} [S]$,

dove il pedice $_{ini}$ ricorda che J è una iniezione.

In quanto struttura insiemistica, e per quanto astratta, Σ deve essere considerata "significante", e ciò presuppone che ogni enunciato di Σ sia a priori "vero" aut (≡ disgiunzione esclusiva) "falso". Ovviamente nulla del genere è pensabile in S, in cui mancano le stesse nozioni di vero e di falso. Tuttavia l'iniezione J di [Σ] in [S] può attribuire all'enunciato formale A = J(α) di S, dove α è il generico enunciato di Σ, un "valore di verità*$_{\mathcal{M}}$" (convenzionale) $\nu^*_{\mathcal{M}}(A)$ uguale a quello (fattuale) di α in Σ, $\nu_\Sigma(\alpha)$, *sotto certe condizioni di consistenza interna*. La $\mathcal{M}$ che compare più sopra come pedice di "verità*" è una abbreviazione per la coppia ⟨Σ,J⟩.

Supporremo innanzitutto che un sf S ed una iniezione J come parzialmente descritti esistano per la data struttura Σ. Ricordiamo poi che se un certo dominio è *iniettato* in un certo insieme, la cardinalità del secondo deve essere non-minore di quella del primo: quindi, per la (1), la cardinalità di [S] deve essere non-minore di quella di [Σ],

(2) $\mathrm{card}([S]) \geq \mathrm{card}([\Sigma])$.

"Proiettando" la (1) sulla classe degli enunciati otteniamo

(1_{E}) $J: \mathsf{E}_\Sigma \to_{ini} \mathsf{E}_S$,

dove E_Σ è l'insieme degli enunciati di Σ e E_S quello degli enunciati formali di S; quindi

(2_{E}) $\mathrm{card}(\mathsf{E}_S) \geq \mathrm{card}(\mathsf{E}_\Sigma)$.

Diamo ora un esempio elementare delle condizioni di consistenza interna su J nominate nel paragrafo precedente. Sia $\mathsf{E}_S° =: J(\mathsf{E}_\Sigma) \subset \mathsf{E}_S$, e siano A e ¬A enunciati di $\mathsf{E}_S°$; richiederemo allora

[8] A rigore la teoria dei modelli (semantici) dovrebbe trattarsi all'interno di una teoria assiomatica degli insiemi, ma di fatto è sufficiente esprimersi all'interno della teoria intuitiva, fermo restando che il discorso può sempre trasferirsi all'interno di una teoria assiomatica.

che $\nu^*_{\mathcal{M}}(\neg A) = \sim \nu^*_{\mathcal{M}}(A)$, dove ~ vuol dire “opposto” (cioè ~ trasforma il “vero$^*_{\mathcal{M}}$” in “falso$^*_{\mathcal{M}}$” e viceversa). Sotto un conveniente insieme di condizioni di questo tipo, basta imporre che la

(3_{at}) $\nu^*_{\mathcal{M}}(J(\alpha_{at})) = \nu_\Sigma(\alpha_{at})$

sia soddisfatta per ogni enunciato atomico α_{at} di E_Σ per assicurare la validità di una simile

(3bis) $\nu^*_{\mathcal{M}}(J(\alpha)) = \nu_\Sigma(\alpha)$

per ogni generico enunciato α di E_Σ (la dimostrazione è per induzione sulla complessità di α).

Valendo la (3bis) per ogni α di E_Σ, diventa dunque possibile parlare “consistentemente” di un valore di verità$^*_{\mathcal{M}}$ degli enunciati formali di $E_S°$, identificandolo con quello delle loro antimmagini sotto J^{-1}, unicamente determinate in E_Σ. Se tutti gli enunciati di un insieme $K \subset E_S°$ sono veri$^*_{\mathcal{M}}$, diremo che $\mathcal{M}$ è un **modello** (semantico) **di** K **in** S.

Si sarà certamente notato che in nessuna delle precedenti definizioni è intervenuta la natura *deduttiva* del sf S, ma soltanto il suo linguaggio (del 1° ordine) L; sarebbe stato cioè altrettanto consistente sostituire nelle precedenti notazioni S con L. I meccanismi deduttivi di S sono peraltro coinvolti in un fondamentale metateorema (™) (non difficile da dimostrare) secondo il quale la verità$^*_{\mathcal{M}}$ è “ereditata” lungo le dimostrazioni in S: vale a dire, «se un enunciato di $E_S°$ è deducibile in S, [9] *e tutti gli assiomi di* S *sono* (in $E_S°$ e) veri$^*_{\mathcal{M}}$, allora quell’enunciato è vero$^*_{\mathcal{M}}$.» Proprio questo è l’aspetto intrigante della situazione, e rende fedelmente conto della “attività deduttiva” del matematico. Vale a dire, il matematico è in Σ, e (se gli è possibile) applica E_Σ in E_S (per qualche S) mediante J, essendosi preventivamente assicurato che gli assiomi di S siano (in $E_S°$ e) veri$^*_{\mathcal{M}}$. Messo di fronte al generico enunciato α di E_Σ, egli lo trasforma in $A =: J(\alpha)$; se è in grado di *dedurre* A come teorema di S, allora in forza di (™) sa che A è vero$^*_{\mathcal{M}}$, e quindi conclude che α è vero (in Σ) *senza averlo verificato fattualmente*.

Naturalmente un tale risultato ha senso soltanto se *non tutti* gli enunciati di E_S (o almeno di $E_S°$) sono deducibili, cioè se il sf S (o almeno il più debole sf S° che si ottiene da S eliminandone gli assiomi che non sono in $E_S°$, a parità di linguaggio) è *coerente*. Infatti, partendo da un enunciato α di E_Σ, e passando alla sua negazione non-α (che è ancora banalmente in E_Σ), si ottengono per essi due valori opposti di ν_Σ; quindi, per la (3bis), due valori opposti di $\nu^*_{\mathcal{M}}$ per $J(\alpha)$ $J(\text{non-}\alpha)$. Ma se tutti gli enunciati di $E_S°$ sono deducibili, in forza di (™) sia $J(\alpha)$ che $J(\text{non-}\alpha)$ (che sono entrambi in $E_S°$) devono essere veri$^*_{\mathcal{M}}$: si ottiene una contraddizione. È pertanto necessario anteporre all’intera argomentazione il vincolo che S, o almeno S°, sia *coerente*. Concludiamo con la seguente importante definizione. Se per una data struttura insiemistica astratta Σ esiste un sf del 1° ordine S e

[9] Beninteso, presupponiamo qui che la deducibilità in S sia quella “naturale”, fondata sull’unica regola cosiddetta “di separazione”: “da A e da A⇒B, inferisci B”.

una iniezione J di $[\Sigma]$ in $[S]$ sotto le descritte condizioni, *e* $S°$ *è coerente*, Σ si dice **formalizzabile in** S **sotto** J, e diventa a tutti gli effetti una **teoria matematica**.

Fino a questo punto, il mondo della realtà empirica non è stato nemmeno nominato. Un passo ulteriore, concettualmente privo di difficoltà, ci porta peraltro alla nozione di (possibile) **formalizzabilità** di una **struttura insiemistica empirica** $\mathcal{E}$. Infatti quest'ultima è del tutto assimilabile alla precedente struttura insiemistica astratta Σ, nel senso che $\mathcal{E}$ è significante esattamente come Σ. Se dunque esistono un sf T del 1° ordine e una iniezione I: $[\mathcal{E}] \to_{ini} [T]$ aventi proprietà rispetto a $\mathcal{E}$ analoghe a quelle di S e J rispetto a Σ, detto $\mathsf{E}_{\mathcal{E}}$ l'insieme degli enunciati (empirici) di $\mathcal{E}$, il rapporto della coppia $M \equiv \langle \mathcal{E}, I \rangle$ con T e quello della coppia $\mathcal{M} \equiv \langle \Sigma, J \rangle$ con S sono completamente analoghi. In particolare, se tutti gli enunciati di un insieme $H \subset \mathsf{E}_T° =: I(\mathsf{E}_{\mathcal{E}}) \subset \mathsf{E}_T$ sono veri$*_M$, diremo che M è un **modello empirico di** H **in** T; sempreché T, o almeno il sf $T°$ che da esso si ottiene eliminandone gli assiomi che non sono in $\mathsf{E}_T°$, sia supposto coerente. [10]

D'altra parte, come per la data struttura astratta Σ non si poteva contare sulla esistenza di un sf S e di una iniezione J di $[\Sigma]$ in $[S]$ con le descritte proprietà, altrettanto non si potrà contare, per la data struttura empirica $\mathcal{E}$, sulla esistenza di un sf T e di una iniezione I di $[\mathcal{E}]$ in $[T]$ con le analoghe proprietà. Anzi, in senso intuitivo/psicologico si percepisce $\mathcal{E}$ come "più lontana" da un possibile T di quanto lo sia Σ da un possibile S. In ogni modo, se per la data $\mathcal{E}$ esistono un T e una I come descritte, la struttura empirica $\mathcal{E}$ potrà dirsi **formalizzabile in** T **sotto** I. Anche se gli stessi protagonisti dell'operazione spesso non ne hanno una totale consapevolezza, la fiducia che questo sia il caso, e la speranza di identificare almeno un T e una I a partire da $\mathcal{E}$, *motivano e governano sostanzialmente l'attività dei fisici matematici e dei fisici teorici*.

È interessante − in verità più da un punto di vista pratico che di principio − considerare la possibilità che la distanza da $\mathcal{E}$ a T sia coperta in due passi successivi, e cioè mediante la composizione ($\circ$) di una prima iniezione G di $[\mathcal{E}]$ in una conveniente $[\Sigma]$, e di una seconda iniezione J di $[\Sigma]$ in un conveniente $[T]$. Esaminiamo la natura di una iniezione come G, che applica una struttura significante come $\mathcal{E}$ in un'altra struttura significante come Σ. Poiché sia in $\mathcal{E}$ che Σ sono definiti valori di verità degli enunciati, diciamo $\nu_{\mathcal{E}}$ e ν_{Σ}, è naturale imporre che sia

(4) $\quad \nu_{\mathcal{E}}(\mathcal{A}) = \nu_{\Sigma}(G(\mathcal{A}))$

[10] L'idea di collegare alla matematica la nozione di "spiegazione scientifica" risale nientemeno che a Platone: nel suo Filebo, egli dà per primo chiara espressione al fatto che una cosiddetta "scienza" è tale nella misura in cui contiene della matematica. Non è quindi illegittima l'opinione dello storico della matematica M. Kline, secondo cui «la realizzazione suprema della cultura greca consistette nel riconoscimento del valore della matematica nella ricerca scientifica.» (in "Mathematical Thought from Ancient to Modern Times", vol II (1972)). L'idea di Platone si ritrova tal quale in Kant: «Una disciplina contiene tanta scienza quanta è la matematica che usa.» (in "Metaphysische Anfangsgründe der Naturwissenschaft" (1786)).

per ogni enunciato (empirico) $\mathcal{A}$ di $\mathsf{E}_{\mathcal{E}}$. Questa uguaglianza implica certe condizioni di autoconsistenza su G, delle quali illustriamo il seguente esempio. Siano $\mathcal{A}$ e $\mathcal{B}$ enunciati arbitrari di $\mathsf{E}_{\mathcal{E}}$: allora "$\mathcal{A}$ e/o $\mathcal{B}$" è manifestamente un enunciato di $\mathsf{E}_{\mathcal{E}}$, e G($\mathcal{A}$ e/o $\mathcal{B}$) è in $\mathsf{E}_{\mathcal{E}}°$ =: $G(\mathsf{E}_{\Sigma}) \subset \mathsf{E}_{\mathcal{E}}$. Risulta perciò, secondo la (4), $\nu_{\mathcal{E}}(\mathcal{A}$ e/o $\mathcal{B}) = \nu_{\Sigma}(G(\mathcal{A}$ e/o$\mathcal{B}))$. Come è noto dalla logica elementare, dando al valore di verità "vero" il valore numerico 0, ed a quello del "falso" il valore numerico 1, il valore di verità di una disgiunzione e/o è uguale al prodotto numerico dei valori di verità dei disgiunti; quindi, da una parte $\nu_{\mathcal{E}}(\mathcal{A}$ e/o $\mathcal{B}) = \nu_{\mathcal{E}}(\mathcal{A})\, \nu_{\mathcal{E}}(\mathcal{B}) = \nu_{\Sigma}(G(\mathcal{A}))\nu_{\Sigma}(G(\mathcal{B}))$ (in forza della (4)) $= \nu_{\Sigma}(G(\mathcal{A})e/oG(\mathcal{B}))$; e dall'altra, ancora per la (4), $\nu_{\mathcal{E}}(\mathcal{A}$ e/o$\mathcal{B}) = \nu_{\Sigma}(G(\mathcal{A}$ e/o$\mathcal{B}))$. Segue che $(G(\mathcal{A})$e/o$G(\mathcal{B}))$ e $G(\mathcal{A}$ e/o $\mathcal{B})$ devono avere lo stesso valore di verità in Σ, ossia che G deve essere tale da trasformare una disgiunzione di enunciati di $\mathsf{E}_{\mathcal{E}}$ nella disgiunzione dei trasformati dei disgiunti, modulo un'equivalenza. Detto diversamente, G deve essere un isomorfismo, modulo un'equivalenza, rispetto alla disgiunzione. Viceversa, se G soddisfa alle richieste di isomorfismo (modulo equivalenze) di cui abbiamo appena dato l'esempio relativo alla disgiunzione, e imponiamo la (4) su enunciati atomici di $\mathsf{E}_{\mathcal{E}}$, allora la stessa (4) vale per enunciati di $\mathsf{E}_{\mathcal{E}}$ generici (la dimostrazione è al solito per induzione sulla complessità dell'enunciato).[11]

Se per data $\mathcal{E}$ esiste Σ ed esiste G: $[\mathcal{E}]\to_{\text{ini}}[\Sigma]$ sotto tutte le condizioni descritte, diremo di avere una **teoria fisica di $\mathcal{E}$ in Σ sotto** G. Ma questa possibilità non ha un vero valore pratico, perché Σ (o piuttosto una sua parte) è una mera "fotografia astratta" di $\mathcal{E}$ che non aggiunge granché alla nostra capacità di prevedere. Il vero interesse nasce invece se per quella struttura astratta Σ esiste un sf T ed una iniezione J: $[\Sigma] \to_{\text{ini}}[T]$ sotto le solite richieste (in particolare che T, o almeno T°, sia coerente): allora Σ è una teoria matematica, e l'iniezione composta I =: J∘G applica direttamente $[\mathcal{E}]$ in [T] soddisfacendo tutte le necessarie condizioni di consistenza. Se dunque abbiamo una teoria fisica di $\mathcal{E}$ (in Σ sotto G) e Σ è formalizzabile (in T sotto J), allora diremo di avere una **teoria fisico-matematica** (tfm) **di $\mathcal{E}$ in** T **sotto** I = J∘G. Se questo è il caso, è evidente perché una tfm di $\mathcal{E}$ (in T sotto I) fornisce una capacità predittiva intorno agli enunciati di $\mathsf{E}_{\mathcal{E}}$: basta ripercorrere l'analoga procedura che andava da Σ a S. Precisamente, il fisico matematico trasforma il generico enunciato $\mathcal{A}$ di $\mathsf{E}_{\mathcal{E}}$ nell'enunciato A = I($\mathcal{A}$) di $\mathsf{E}_T°$; se sa provare che A è deducibile in T°, (i cui assiomi sono per ipotesi (in $\mathsf{E}_T°$ e) veri$*_M$), allora sa anche che A è vero$*_M$ in forza di (™). In definitiva egli può concludere che $\mathcal{A}$ è vero (in $\mathcal{E}$) *senza* averlo osservato.[12]

[11] Le richieste di isomorfismo modulo equivalenza sono essenzialmente due, quella rispetto alla negazione (G(non-$\mathcal{A}$) e non-G($\mathcal{A}$) hanno lo stesso valore di verità in Σ per ogni $\mathcal{A}$ di $\mathsf{E}_{\mathcal{E}}$), e quella rispetto alla disgiunzione illustrata nel testo, possibilmente estesa, con certe precauzioni, alla disgiunzione *transfinita* (che è essenzialmente il quantificatore esistenziale) se card($\mathsf{E}_{\mathcal{E}}$) non è finito.

[12] La ricognizione degli enunciati empirici di $\mathcal{E}$ deve svolgersi secondo modalità operative inambiguamente convenute, ed è di regola affetta da un certo grado di *imprecisione* diagnostica. Se dunque la nozione di verità empirica si

Una **legge** sulla struttura empirica $\mathcal{E}$ (o **legge empirica**) si esprime tipicamente affermando che un certo predicato $\mathcal{A}$ di $\mathsf{E}_{\mathcal{E}}$ su individui x variabili su un universo $\mathcal{D}$ (non vuoto) [13] ha valore di verità "vero" in $\mathcal{E}$, quindi nella forma

(5) $\quad \nu_{\mathcal{E}}(\mathcal{A}(x)) = 0.$

Poiché per ipotesi x appartiene a $\mathcal{D}$, $\mathcal{A}$ è **collettivizzante** in x ("principio di separazione"), e quindi possiede una **estensione** unicamente determinata $\mathcal{X}$ (quella che formalmente si denota con $\{x|(x\in\mathcal{D})\wedge\mathcal{A}(x)\}$, inclusa in $\mathcal{D}$) per la quale $\mathcal{A}(x)$ equivale a "x appartiene a $\mathcal{X}$". La possibile determinazione di questa estensione $\mathcal{X}$ presenta particolare interesse. Se è disponibile una iniezione G: $[\mathcal{E}] \to_{\text{ini}} [\Sigma]$, dove Σ è la solita struttura insiemistica astratta, e $\mathcal{A}(x)$ si trasforma sotto G in $\alpha(\xi)$ (dove $\alpha = \mathrm{G}(\mathcal{A})$, e $\xi = \mathrm{G}(x)$ è variabile in $\Delta =: \mathrm{G}(\mathcal{D})$), allora la (5) ci porta al predicato su ξ

(5bis) $\quad \nu_{\Sigma}(\alpha(\xi)) = 0.$

Ovviamente anche α è collettivizzante in ξ, ed ha una estensione unicamente determinata $\Xi = \mathrm{G}(\mathcal{X})$ inclusa in Δ (non vuoto), e per la quale $\alpha(\xi)$ equivale a "ξ appartiene a Ξ". La determinazione di Ξ è un problema ormai interno alla struttura astratta Σ, risolto (se possibile) il quale si risale a $\mathcal{X}$ come $\mathrm{G}^{-1}(\Xi)$. Ma evidentemente il caso di vero interesse è quello in cui Σ è *formalizzabile* in un sf del 1° ordine T (per il quale il solito T° è supposto coerente) sotto J, e quindi Σ è una teoria matematica; allora abbiamo una genuina tfm di $\mathcal{E}$, e con essa una capacità predittiva intorno a $\mathcal{E}$ fondata sulla legge empirica (5). Probabilmente il lettore si è ormai familiarizzato con i meccanismi che presiedono a questa capacità. Se ancora $\mathrm{I} =: \mathrm{J}\circ\mathrm{G}$, posto $\mathrm{A} =: \mathrm{I}(\mathcal{A})$ e $\mathrm{x} =: \mathrm{I}(x)$, la (5) si trasforma nel predicato su x

(5ter) $\quad \nu_{\mathrm{M}^*}(\mathrm{A}(\mathrm{x})) = 0,$

e la determinazione di $\mathcal{X}$ è un problema "matematico", interno al sf T°. Questo problema prende usualmente la forma di una **equazione nella incognita** x sotto la (debole, ma sufficiente)

---

accompagna ad una inevitabile imprecisione (o meglio ad una imprecisione relativa), almeno un limite superiore di questa imprecisione relativa, diciamo $\varepsilon > 0$, deve comparire in tutti i ragionamenti che coinvolgono la validazione o invalidazione degli enunciati empirici. Ciò è quanto *avviene normalmente* nell'attività scientifica esatta, anche se ε non è menzionata esplicitamente. Volendolo evitare, a rigore dovremmo inquadrare il nostro discorso in una logica più generale di quella naturale, e che è stata detta "paracoerente": in essa, la non-verità di un enunciato non implica in generale la verità della sua negazione, e le verità separate di due enunciati non implicano in generale la verità della loro congiunzione. A nostra conoscenza, questa opzione è del tutto ignorata nella pratica dell'attività scientifica-esatta. L'ovvia alternativa è quella di sbarazzarsi di ε semplicemente immaginando di operare in un mondo osservabile *ideale* in cui sia legittimo fare $\varepsilon \to 0$, come di norma si sottintende. Questa possibilità sussiste anche in meccanica quantistica, ove si sposti l'attenzione dal singolo sistema, sul quale $\varepsilon = 0$ non è attingibile a causa delle relazioni di incertezza, ad una folla (o ensemble) di sistemi identici (cfr. il teorema di Feenberg: «la funzione d'onda di un sistema è unicamente determinata con precisione *arbitraria*, a meno di un fattore di fase *costante*, misurando la densità di probabilità della folla e la sua derivata temporale»). Non discutiamo qui questo tema, un'analisi approfondita del quale ha portato alla nota proposta di una **logica quantistica** (QL, Quantum Logic), perché limitiamo la nostra attenzione alle tfm macroscopiche o classiche.

[13] Questo predicato può essere di tipo atomico $(n\geq 0)$-ario su un prodotto cartesiano di grado n di differenti universi, oppure risultare da composizioni logiche di predicati di questo tipo, ecc.

condizione che T° sia un sf (del 1° ordine e) **egualitario** [14]. Per definizione, un equazione in un sf egualitario consiste nella eguaglianza di due suoi termini V e W in almeno uno dei quali figura x, quindi nella

(6) $V(x) = W(x)$;

una sua **soluzione** è un individuo $\underline{x}$ di $D =: I(\mathcal{D})$ (non vuoto) per il quale l'enunciato $V(\underline{x}) = W(\underline{x})$ è deducibile in T°. L'insieme-estensione X della (6), incluso in D, è l'insieme di tutte e sole le sue soluzioni, e $\mathcal{X} = \Gamma^{-1}(X)$. Se $X = \varnothing$, la (6) non ha soluzioni, e il modello equazionale che la genera si dice **malsano**. Parecchi modelli equazionali malsani (anche se essi sembrano sanissimi ad un esame superficiale) sono o dovrebbero essere ben familiari al fisico matematico: ad esempio, sotto la maggior parte delle (ragionevoli) condizioni accessorie, il modello idrodinamico ideale di Eulero è malsano. Se invece X = D, la (6) ha "troppe" soluzioni, e non ci dà in pratica alcuna nuova informazione; in questo caso, raro e banale, il modello da cui essa deriva si dice **vuoto**. La situazione di maggior interesse è ovviamente quello in cui X consta di unico elemento ("X è un singoletto"), ovvero in cui la (6) ha una e una sola soluzione. In generale, cioè anche quando non si traduce in una *equazione* di T°, una legge empirica (5) tale che esista esattamente un individuo $\underline{x}$ di $\mathcal{D}$ per cui essa è vera in $\mathcal{E}$, si dice **deterministica** in x. A questa situazione fa riscontro un predicato matematico del tipo (5ter) [del tipo (6)] per il quale esiste uno e un solo individuo $\underline{x}$ di D che lo soddisfa [una e una sola soluzione $\underline{x}$ in D].

Ci si aspetta che, essendo significante, una teoria matematica debba risultare sia coerente che completa. Possiamo rinunciare ad almeno una di queste proprietà? Come abbiamo visto, una teoria matematica deve essere "fondata", cioè deve essere formalizzabile in un sf T nel senso che abbiamo appena illustrato. Ora in forza del teorema (™) E_T deve essere coerente in T, o almeno $\mathsf{E}_T°$ deve essere coerente in T°. Invece T° può ben essere incompleto: se un enunciato α di E_Σ va in un $A = J(\alpha)$ di $\mathsf{E}_T°$ che non è né deducibile né refutabile in T°, non avremo alcuna informazione sulla verità (in Σ) di α o di non-α attraverso T°, ma da questo fatto *non nascono contraddizioni.* In conclusione possiamo accettare teorie matematiche fondate in un sf T° *incompleto* (quindi teorie incomplete) ma in nessun caso in un sf T° *incoerente* (quindi teorie incoerenti).

D'altra parte T (o T°) deve essere abbastanza ricco per permetterci di rappresentarvi una teoria matematica Σ di qualche interesse, e soprattutto di interesse fisico-matematico. Secondo un diffuso e giustificato convincimento, la base canonica da cui partire per costruire teorie matematiche, e in particolare le teorie matematiche tipicamente associate alle tfm, è la teoria degli

[14] Cioè che T° annoveri tra i suoi assiomi gli assiomi dell'uguaglianza (oltre agli assiomi logici e di quantificazione che ne specificano il carattere di sf del 1° ordine, logico-quantificato).

insiemi. [15] Teorie matematiche come la Teoria dei Gruppi, l'Aritmetica dei Numeri Naturali o Reali o Complessi, la Topologia, l'Analisi, l'Analisi Funzionale, .. ecc., continuamente usate dai fisici matematici e teorici, devono cioè sostanzialmente vedersi come proiezioni significanti, o modelli semantici, della sopraddetta teoria degli insiemi (o di sue parti) in una delle sue assiomatizzazioni correnti, [16] con l'aggiunta di opportuni **assiomi specifici** e connesse **definizioni specifiche**.

È chiaro che sia la coerenza che la completezza di un sf T del 1° ordine sono sue proprietà sintattiche, che esso può a priori avere o non avere; ma se T ha un modello $\mathcal{M} \equiv \langle \Sigma, J \rangle$, allora esso (o almeno la sua parte T° definita al solito modo in termini di $\mathcal{M}$) deve essere coerente. Se T non è coerente, la sola possibilità è che non abbia alcun modello. E stato presto dimostrato (E. Post, 1921) che un **Calcolo Enunciativo** CE ha modelli, ed è tale che (i) i suoi Teoremi ($\equiv$ "assiomi o teoremi") sono veri in un suo qualunque modello, e che (ii) i suoi enunciati veri in un suo qualunque modello sono suoi Teoremi. Molto più laboriosamente, un analogo asserto è stato ottenuto da K. Gödel (1930) per un **Calcolo Predicativo del** 1° **ordine**. Come si intuisce, la parte difficile di questi teoremi è la seconda (ii); e anzi, è proprio alle loro seconde parti che ci si riferisce usualmente quando li si nominano.

Mediante strumenti di legittimità inquestionabile, nel 1931 Gödel [17] ha dimostrato che se un sf abbastanza ricco (in pratica, capace di descrivere l'aritmetica dei naturali) e "effettivamente" descrivibile, quindi in particolare finitamente assiomatizzabile ($\equiv$ con numero di assiomi e schemi di assiomi, che sono insiemi di assiomi con comune struttura, finito), è coerente, allora deve essere incompleto; e quindi, per trasposizione, che se esso è completo, deve essere incoerente. Inoltre, la necessaria incompletezza di un sf $\mathcal{T}$ coerente e abbastanza ricco è di tipo "essenziale": vale a dire, se A è un generico suo enunciato non deducibile né refutabile, e A stesso (o $\neg$A) viene aggiunto come assioma agli assiomi di $\mathcal{T}$, il risultante sf (certamente coerente se $\mathcal{T}$ lo era) è ancora, si dimostra, incompleto.

Compreso nello stesso lavoro di Gödel è un fondamentale corollario di questo risultato, il quale afferma che la possibile coerenza del sf considerato non può dimostrarsi mediante la metamatematica che *esso stesso* rende disponibile (o se si preferisce, mediante metodi

[15] «È ragionevole pensare che la totalità della Matematica possa essere derivata da una sola sorgente: la teoria degli insiemi.» (N. Bourbaki).

[16] Ad esempio la ZFC (assiomatizzazione di Zermelo-Fraenkel + assioma di scelta (Choice)). Se effettivamente utilizzata, questa inclusione assicura con enorme larghezza il carattere *ramificato* e *proliferante* dei teoremi della teoria, e quindi la non-banalità della associata tfm.

[17] K.Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I" Monatschefte für Mathematik und Physik vol 38, 173-198 (1931). Di questo epocale lavoro esistono traduzioni (e anche *diverse* traduzioni) praticamente in tutte le lingue, compreso l'italiano (in E. Agazzi, "Introduzione ai problemi dell'assiomatica", Vita e Pensiero (1962)), e naturalmente l'inglese ("On Formally Undecidable of «Principia Mathematica» and Related Systems", Basic Books (1965), ristampato da Dover; anche in J. van Heijenoort, "A Source Book in Mathematical Logic, 1879-1931", Harvard Un. Press (1967), e anche in K. Gödel, "Collected Works" (a cura di S. Feferman), vol I, Pubblications 1929-1936, Oxford Un. Press (1986).

formalizzabili all'interno di quel sf); vale a dire, mediante la più ampia metamatematica *leale*. Una eventualità di questo tipo (che è stata pittorescamente paragonata a quella di «tirarsi fuori dalla palude afferrandosi per il proprio codino», alla maniera del Barone di Münchausen), è infatti manifestamente la più permissiva entro la quale una prova di coerenza si configurerebbe ancora come del tutto leale. Dobbiamo dunque accettare di convivere al più con la *fiducia*, e non con la *certezza* (ottenuta con mezzi leali), che un sf abbastanza ricco sia coerente: una teoria matematica che sia espressione semantica di un sf $\mathcal{T}$ semplicemente *supposto* coerente potrebbe in realtà portare a contraddizioni, e può essere usata soltanto con il rischio che ciò possa avvenire da un momento all'altro. Insomma la coerenza di un sf $\mathcal{T}$ di qualche interesse non può dimostrarsi "al suo interno": può invece (con la dovuta dose di ingegno) dimostrarsi "dall'esterno", cioè facendo uso allo scopo (in senso metamatematico) di asserti esterni a $\mathcal{T}$, o addirittura di un sf più forte di $\mathcal{T}$, a sua volta supposto coerente. Ma oltre a costituire un gioco *sleale*, una tale possibilità dà origine ad un rinvio infinito ad altri sf. L'unica via di uscita da questo dilemma sembra dunque quella di (i) rassegnarsi ad usare sf *potenzialmente* incoerenti, oppure (ii) limitarsi a considerare sf così poveri che la prova di una loro (possibile) coerenza sia ottenibile mediante una metamatematica elementare, sulla cui legittimità non vi sia alcun modo di eccepire. Casi del genere non mancano: l'esempio più naturale è quello di CE, che si può appunto dimostrare coerente, completo e decidibile sulla base di mezzi metamatematici elementari. Ma è ovvio che l'opzione (ii) è troppo restrittiva.

A questa delicatissima impasse, [18] nel corso del XX secolo si sono sovrapposti certi (meta)teoremi cosiddetti "di coerenza relativa" e "di indipendenza", che hanno ulteriormente scosso le primitive speranze dei formalisti di pervenire a prove "interne" di coerenza dei sf di interesse. In particolare, nel 1940 Gödel (ancora lui!) ha provato [19] che, supponendo la teoria degli insiemi secondo ZF coerente, essa rimane coerente aggiungendovi come assioma l'assioma di scelta (AC), oppure l'ipotesi generalizzata del continuo (IGC). Più in generale, nel 1963 P. Cohen ha dimostrato [20] l'indipendenza di AC e di IGC dagli assiomi di ZF, e contemporaneamente l'implicazione IGC $\Rightarrow$ AC. [21] Emerge così una *pluralità* di matematiche altrettanto legittime dal punto di vista

[18] Sapidamente sintetizzata da H. Weyl nell'aforisma «Dio esiste perché la matematica è certamente non contraddittoria; ma anche il diavolo esiste, perché non possiamo dimostrarlo.»
[19] K. Gödel, "The consistency of the axiom of choice and of the generalized continuum hypotesis with the axioms of set theory", Princeton Un. Pr. (1940). Questa breve monografia, tratta da lezioni tenute da Gödel a Princeton, era stata preceduta da alcune note sullo stesso tema sin dal 1938.
[20] I risultati di Cohen sono ordinatamente esposti e ampliati nella sua monografia "Set Theory and Continuum Hypotesis", Benjamin (1966).
[21] Ricordiamo in cosa consistono AC e IGC. (1) AC: "per una qualunque famiglia di insiemi non vuoti e disgiunti, esiste un insieme che ha esattamente un elemento in comune con ciascun insieme della famiglia"; (2) IGC: "Per ogni $n \geq 0$, non esistono cardinalità comprese tra $\aleph_n$ e $\aleph_{n+1}$". Le cardinalità degli $\aleph$ si definiscono ricorsivamente come segue: a) $\aleph_{n+1}$ è la cardinalità dell'insieme delle parti (vedi A.3) di un insieme di cardinalità $\aleph_n$; b) $\aleph_0$ è la cardinalità dell'insieme dei numeri naturali. Per $n = 0$, IGC si dice semplicemente "ipotesi del continuo" (IC), e risale a G. Cantor.

logico (ma ovviamente altrettanto esposte al rischio di essere incoerenti), tra le quali è assai problematico (o imbarazzante) optare.

Queste sconcertanti conclusioni, che il già citato M. Kline non esita a definire come «la perdita della certezza matematica» [22] , sono alla base di una intera letteratura di varia estrazione e momento, che va dal catastrofismo senza appello alla reazione dogmatica della scuola intuizionista (secondo gli intuizionisti, da una parte la matematica è coerente perché lo garantisce il suo significato intuitivo, e dall'altra sia AC che IGC sono inaccettabili per l'opposta ragione). Tra questi estremi, il compromesso più equilibrato e pragmatico sembra essere quello oggi dominante nel mondo dei cosiddetti "matematici fattivi": se da un lato la coerenza di una teoria matematica di qualche interesse non si può dimostrare con mezzi leali, dall'altro essa si *constata* giorno dopo giorno nella assenza di contraddizioni che potrebbero nascervi; e quanto alla incompletezza che automaticamente consegue dal teorema di Gödel, è possibile accettarla e convivere con essa. [23] Insomma il matematico fattivo continua con imperturbabile operosità ad allargare i confini di un suo mondo già incredibilmente vasto, e l'ultima preoccupazione che lo tocca è l'idea di muoversi su un terreno insicuro, che potrebbe inghiottirlo da un momento all'altro. Semplicemente, ha fiducia che questo non avvenga; ma se dovesse sbagliarsi, è pronto a (tentare di) porre rimedio alle nuove possibili difficoltà. E non gli si può dare torto: occorre continuamente ripetersi che la matematica non è altro che un prodotto, certamente tra i più alti, della immaginazione, della inventività, della curiosità e della tenacia dell'uomo, e proprio per questo la sua rotta va continuamente sorvegliata e corretta, facendo tesoro dell'esperienza passata e delle possibili divergenze di opinione che nascano tra gli addetti ai lavori. [24]

Superfluo aggiungere che la crisi fondazionale che investe l'intera matematica di qualche importanza si rifletta praticamente su tutte le tfm; ma anche in questo caso, la situazione è di per sé ben lontana dal produrre effetti negativi concreti. Infatti il fisico teorico "medio" sembra sostanzialmente disinteressato alla minacciosa divaricazione presente tra le (presunte) teorie matematiche che usa nel suo lavoro quotidiano e i sf che dovrebbero "fondarle": per fortuna, egli si comporta come se ne ignorasse l'esistenza (e spesso la ignora di fatto, almeno nella sua accezione più profonda e nelle sue implicazioni più remote). Come riconosce lo stesso A. Einstein, dopotutto il fisico teorico ha un rapporto «piuttosto spregiudicato» con la matematica: usa di volta in volta

[22] M. Kline, "Mathematics: The Loss of Certainty", Oxford Un. Press, (1980). Ovviamente il senso del titolo del libro è quello della perdita di una certezza vivamente "vagheggiata". Il corollario del teorema di incompletezza di Gödel cadde come una scure su questo sogno.

[23] Dopotutto non sembra un gran sacrificio accettare che certi presunti enunciati di quella teoria (come ad es. la congettura di Goldbach nella teoria dei naturali − ogni numero pari > 2 è somma di almeno una coppia di primi), di fatto inevitabilmente veri o falsi, forse non siano né provabili né refutabili con mezzi leali (o almeno non lo siano stati fino ad oggi).

[24] «Per venticinque secoli i matematici hanno continuamente corretto i loro errori, e come conseguenza hanno visto la loro disciplina arricchirsi, non impoverirsi; ciò dà loro il diritto di guardare al futuro con animo sereno.» (N. Bourbaki).

quella che di più gli conviene per meglio conseguire il suo vero fine, che è quello di disporre di una capacità predittiva sempre più estesa ed efficace nei confronti del mondo fenomenico.

Resta da chiedersi (ma la domanda è a questo punto abbondantemente retorica) che cosa spinga i logici a ricercare la formalizzazione in senso stretto delle teorie matematiche. Il punto è che il livello dei segni di un alfabeto elementare (segni sulla carta, spots magnetizzati su un hard disc, ecc.) di un sf è l'ultimo al di là del quale *non è più possibile arretrare*: esso offre un *massimo di concretezza*, quella tipica dell'attività di un automa. [25] Via formalizzazione stretta, un tale automa potrebbe *enumerarci* buona parte delle *certezze* (non tutte) relative alla matematica associata alle nostre tfm. Ma il perseguimento di un tale obiettivo, e con esso quello di ridurre la "spiegazione" ultima dei meccanismi che governano il mondo fenomenico (o una sua sezione) alla esecuzione di convenienti *algoritmi aperti* [26] (praticamente tutte le teorie matematiche di qualche interesse sono indecidibili), in linea di principio è destinato a rimanere sotto la minaccia della comparsa di contraddizioni inaspettate. Per concludere, va infine ricordato che anche la moderna "Computer Science" sembra rimuovere questo tipo di minaccia, concentrando pragmaticamente buona parte della sua attenzione sull'efficienza e l'affidabilità degli algoritmi (aperti o chiusi che siano) e sui vantaggi del loro impiego in termini di rapporto costi/benefici.

Riteniamo sufficiente alle nostre finalità chiudere qui la nostra "libera introduzione" al **problema epistemologico** (il lettore interessato ad approfondimenti potrà sempre ricorrere alla letteratura specializzata, di fatto non particolarmente vasta). Alla luce di quanto abbiamo illustrato, dovrebbe comunque essere raggiunto il nostro presente obiettivo didattico: l'essere chiaro cioè *cosa siano le tfm* e perché esse si dicano anche "teorie fisiche formalizzate" (oppure, "formalizzabili"), e infine, in che senso esse coincidano con le scienze esatte del linguaggio corrente.

[25] Naturalmente qui si vorrebbe una definizione precisa di cosa debba intendersi per "attività di un automa". Formulata in questi termini, la questione non sarebbe tuttavia ben-posta. Rinviamo il lettore curioso su questo punto alla cosiddetta "tesi di Church", che risale al 1936. La tesi di Church si collega al teorema secondo il quale le funzioni cosiddette "λ-definibili" di Kleene, le funzioni cosiddette "generalmente ricorsive" di Gödel, e le funzioni cosiddette "Turing-computabili" sono le *stesse* funzioni. Secondo la tesi di Church, un automa è un qualsiasi marchingegno capace di calcolare i valori di questo tipo di funzioni, *ma niente di più complicato*.

[26] Gli algoritmi che affidiamo al nostro computer sono tutti "chiusi" nel senso che il computer si arresta ad un certo punto della loro esecuzione. Tuttavia l'arresto è di solito dovuto ad un ordine incorporato nel programma di esecuzione, *e il criterio che fissa tale ordine è di solito arbitrario*, cioè dettato da un giudizio di solito opinabile di chi ha compilato il programma. Soltanto se quest'ordine di arresto "artificiale" manca, essendo sostituito da un ordine "naturale" che

I.2) RAGIONAMENTO DEDUTTIVO E RAGIONAMENTO INDUTTIVO

Da una teoria fisica formalizzata, lo studioso dei fenomeni naturali, il “fisico” nell’accezione dell’etimo, si aspetta che attribuendo ai suoi termini i loro significati intuitivi nel linguaggio naturale, e analogamente traducendo i suoi enunciati, si ottengano i corrispondenti oggetti empirici, operativamente ben definiti; e in particolare riferendoci ai suoi assiomi, e ai teoremi che da essi conseguono attraverso le relative regole di inferenza (di larga massima, quella della deduzione naturale), si ottengano verità empiriche verificate o verificabili entro i limiti delle approssimazioni osservative, oppure nel limite ideale in cui la massima imprecisione relativa tende a zero.

Così come è stata definita, una tfm risulta **affidabile** (o **presuntivamente corretta**) in quanto ogni Teorema del dominio dell’interpretazione naturale sia convalidato dall’esperienza; e complementarmente, risulta **adeguata** (o **presuntivamente completa**), rispetto all’ambito (o dominio) empirico preso in esame, in quanto ad ogni verità osservata in quell’ambito corrisponda un Teorema del dominio dell’interpretazione naturale. È chiaro, tuttavia, che la presunta correttezza/completezza della tfm, nella interpretazione naturale, risulta continuamente esposta alla prospettiva di essere falsificata dal raffinarsi delle nostra capacità osservative/transinduttive, e in secondo ordine deduttive. Se questo si verifica, anche in un solo caso o classe di casi indiscutibilmente accertati, la tfm dovrà essere respinta, e se possibile sostituita con un’altra adeguata alla nuova situazione. Va anche detto, a questo proposito, che l’eventualità che una tfm consolidata diventi *totalmente inutilizzabile* a seguito della falsificazione di qualche sua conclusione è piuttosto improbabile: come infatti la storia delle scienze esatte ci insegna, nei casi più noti ed importanti sono bastate *modifiche* − in fattispecie generalizzazioni, sebbene *sostanzialmente non banali* −, alla luce delle quali la vecchia tfm è risultata compatibile con la nuova in quanto identica ad un particolare caso-limite di quest’ultima, continuando a essere “praticamente” corretta in condizioni prossime alle condizioni-limite. In altre parole, e come meglio vedremo più avanti, si è autorizzati a pensare che di solito la nuova teoria “raffini” la vecchia piuttosto che negarla radicalmente. Questo fatto è talvolta proposto come **principio del carattere cumulativo** delle teorie scientifiche esatte.

Su un altro versante, dobbiamo ricordare che una parte preponderante delle nostre attività mentali è legata al ragionamento induttivo. Con un inevitabile grado di approssimazione, questo attributo qualifica quel tipo di attività essenzialmente intuitiva/immaginativa che partendo da asserti isolati e particolari tenta di pervenire (o perviene) a conclusioni *che comunque li superano*. Se poi

---

dice “arrestati perché hai realmente *concluso* il tuo lavoro”, e il computer si ferma, siamo di fronte ad un algoritmo genuinamente chiuso.

tali conclusioni giungono ad affermare una universalizzazione contestuale degli asserti di partenza, cioè equivalgono a leggi generali all'interno di un dato contesto, si parla di **ragionamento induttivo in senso stretto**. In pratica, un ragionamento induttivo rende plausibile (con vario grado di "credibilità", o «probabilità induttiva» (R. Carnap)) una certa conclusione, senza genuinamente implicarla. I tentativi di razionalizzare e formalizzare il ragionamento induttivo non sembrano aver condotto finora a conclusioni definitive, al punto che da parte di molti si contesta l'esistenza di una vera e propria "logica induttiva". Ciò non toglie che il ragionamento induttivo abbia un'importanza capitale, perché senza di esso nemmeno la logica dei sistemi formali potrebbe esistere, sia di per sé che come strumento di conoscenza "esatta" del mondo fenomenico; e ovviamente, nemmeno potrebbe esistere lo stesso universo astratto delle matematiche.

È anche chiaro che la scienza basa la sua fiducia nell'induzione sul successo che essa le ha fin qui assicurato, cioè *giustifica l'induzione mediante un'altra induzione*. Questa non profondissima scoperta di segno scettico, con le pretese difficoltà che ne deriverebbero, è quello che i filosofi chiamano "circolo vizioso di D. Hume". Essenzialmente la stessa idea è stata ripresa e portata alle sue estreme conseguenze da K. Popper e dalla sua scuola. [27] Fortunatamente per le scienze esatte, disponiamo tuttavia di amplissimi sistemi di verità empiriche "previste teoricamente", che lasciano ben poco spazio allo scetticismo di Hume-Popper; senza contare che esso è *per un verso ineludibile e per l'altro sterile*. [28]

[27] Come è noto, Popper rifiuta di ammettere che un sistema di conoscenze avente per oggetto il mondo fisico possa essere comunque convalidato dall'esperienza, e quindi rifiuta il metodo induttivo come base fondazionale delle scienze empiriche. Secondo Popper ("Logik der Forschung", Springer (1934); 1[a] trad. ingl., "Logic of Scientific Discovery", Hutchinson (1959); trad. it., "Logica della scoperta scientifica", Einaudi (1970), § 4-6 e § 21-22), un asserto T di quel sistema di conoscenze S ha comunque la natura di un quantificatore universale ($T \equiv \forall(x)P(x)$) su un dominio *non finito*, e quindi non può essere convalidato da un numero necessariamente finito di verifiche. Al più, S potrà essere «corroborato» da conferme sperimentali; e da ciò deriva la provvisorietà e l'incertezza del sapere scientifico, nonché l'impossibilità che esso aspiri a raggiungere verità definitive. Al contrario, una sola istanza negativa x_0 basta ad invalidare T, ammesso e non concesso di poter refutare x_0 isolatamente, e non in congiunzione con altre x_1, x_2, … (Duhem). Nelle notazioni formali standard (qui usate per brevità), ciò avviene perché $\neg P(x_0) \Rightarrow \exists(x)\{\neg P(x)\} \Leftrightarrow \Leftrightarrow \neg\forall(x)P(x)) \equiv \neg T$; il che induce Popper a proporre come «criterio di demarcazione» tra la "scientificità" e la "non-scientificità" di S la sua «falsificabilità» piuttosto che la sua «verificabilità». Vale a dire, S è falsificabile, e quindi scientifico, se esiste qualche suo enunciato (un suo «falsificatore potenziale») che è in contraddizione con S; ed è «falsificato» se tale falsificatore (che secondo Popper (loc. cit. § 22) sarà «di bassa universalità») è convalidato dall'esperienza. Confessiamo un certo imbarazzo di fronte a queste tesi (già presenti embrionalmente in Bacone (Novum Organum", 1620), che pure hanno avuto un ruolo centrale nella gnoseologia popperiana: la tentazione di associarle al celebre giudizio di C.F. Gauss sulla distinzione kantiana tra proposizioni analitiche e sintetiche è veramente forte. Ciò che più colpisce, agli occhi dell'uomo di scienza "normalmente" estraneo al mondo della filosofia, è la prolissità (rispetto agli standard logici dell'epoca − siamo nel 1934!) con cui esse tesi sono esposte, *nonché la loro totale ovvietà*; e in senso opposto, le innumerevoli discussioni cui hanno dato vita nella letteratura epistemologica degli ultimi 60/70 anni. Quasi superfluo aggiungere che le soprariferite idee di Popper hanno avuto un impatto trascurabile sulla grande maggioranza degli scienziati e degli stessi logici, e (ovviamente) sulle loro concrete attività. Dovrebbe esser chiaro, infine, che una tfm come qui definita supera il criterio di demarcazione popperiano.

[28] Una semplice metafora può meglio illustrare la situazione. Sotto molti aspetti, l'obiettivo di una tfm può paragonarsi a quello di decriptare una successione potenzialmente infinita di messaggi cifrati. È allora ovvio che non potremo mai esser certi che il messaggio successivo sarà decriptato come i precedenti (pur nella confidenza che la chiave di codifica non sia cambiata); ma questo fatto non può e non deve tradursi in un dubbio di principio sulla legittimità ed efficacia dell'uso del decodificatore. Un'altra nota metafora, più o meno equivalente sotto il presente punto di vista, paragona l'attività scientifica esatta al tentativo di comprendere le regole di un gioco competitivo leale − ad esempio, non

La stessa ripetibilità *indefinita* della capacità predittiva di una tfm affidabile resta inoltre sul piano della *confidenza* (per quanto alta questa possa essere), in quanto un oggetto fisico radicalmente caotico non sarebbe prevedibile da alcuna tfm; diversamente detto, una rappresentazione predittiva dell'oggetto fisico considerato si deve comunque fondare sulla presunzione di una sufficiente *regolarità* di quest'ultimo. [29] A ben pensarci, è proprio la ricerca di queste regolarità (supposto che esistano) il più grande incentivo che ci sollecita alla scienza (non soltanto in senso stretto), e il loro scoprirle/verificarle/prevederle la più grande gratificazione che ne deriviamo; al punto che si potrebbe ragionevolmente proporre l'equivalenza (informale) "ricerca scientifica" ≡ "ricerca di regolarità non-banali del mondo fenomenico". "Verificare le regolarità del mondo" significa provare che da esperimenti identici si ricavano (o che situazioni identiche generano) risultati identici; ma è importante specificare in modo preciso cosa si debba intendere per gli uni (le une) e per gli altri. I non infrequenti casi in cui esperimenti/situazioni *quasi* identici/identiche danno luogo a risultati *molto* diversi, e diversi in modo disordinato e imprevedibile, sono detti "critici" o "singolari". Tuttavia le tfm singolari prevedono il carattere caotico dei risultati, e talvolta ne colgono certi possibili aspetti regolari. Si aggiunge dunque una terza proprietà da richiedere ad una tfm (oltre alla sua affidabilità/adeguatezza): quella della **continuità** (o come più spesso si dice, **stabilità**) **strutturale**. La stabilità strutturale di una tfm è per lo più assai difficile da dimostrare, e si è così ancora costretti a convivere con la semplice fiducia che essa sia verificata, più o meno come si convive con la fiducia che la teoria degli insiemi sia coerente. D'altra parte la nota proposta di R. Thom di accettare la stabilità strutturale di una tfm come un suo *assioma* specifico induce non poche perplessità.

Può essere interessante un esame informale della possibilità di emendare "localmente" (o per così dire mediante interventi circoscritti) una teoria scientifica falsificata. È abbastanza chiaro

---

contando le differenze di complessità e di accessibilità, del gioco degli scacchi – assistendo ad un numero potenzialmente infinito di partite. Ci si può allora sempre aspettare che la partita successiva si riveli incomprensibile alla luce delle regole fino a quel momento presunte valide.

[29] A proposito di questa presunzione o fiducia, cioè della connessione che sembra legare la conoscenza del mondo empirico alla matematica, si sono delineate posizioni filosofiche che la interpretano o la giustificano in modi diversi. Ad esempio qualcuno pensa che l'uomo vede ciò che in realtà sta cercando, ovvero che la scienza è solo parzialmente basata sulla realtà sperimentale: la nostra mente sarebbe cioè fatta in modo che gran parte di ciò che vediamo lo vediamo attraverso certe "lenti" (metaforiche) che inconsapevolmente indossiamo. Altri pensano che di fatto l'uomo crea e seleziona esattamente quella matematica che "si adatta" al mondo empirico. Altri ancora ricordano (per la verità piuttosto banalmente) che la matematica offre chiavi di comprensione soltanto di una piccola parte della nostra esperienza, per cui è arbitrario affermare che il mondo nella sua interezza si possa spiegare mediante la matematica. Infine è stata recentemente ripresa una proposta più radicale (ma di trasparente ascendenza pitagorica) secondo la quale la fisica è così ben descritta dalla matematica semplicemente «perché il mondo fisico è completamente matematico, isomorfo ad una struttura matematica che l'uomo viene scoprendo a poco a poco, bit per bit.» (M. Tegmark, 2007). Impossibile non ricordare qui il famoso articolo di E. Wigner "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" (1960) che è all'origine di buona parte delle discussioni sul tema. Nelle parole di Wigner (loc. cit.): «It is difficult to avoid the impression that a miracle confronts us here, quite comparable in its striking nature to the miracle that the human mind can string a thousand arguments together without getting itself into contradictions, or to the two miracles of laws of nature and of the human mind's capacity to divine them.» Si rilegga anche la nota ([10]) della precedente sezione.

che una tale possibilità debba essere esclusa in generale per una tfm formalizzata $\Phi =: \langle T,\mathcal{E},I\rangle$ come qui definita. Se infatti la negazione di un teorema di T è riconosciuta come empiricamente vera, e se T è "logico", per ristabilire la affidabilità/adeguatezza di Φ occorrerà modificare gli assiomi specifici di T (rimuovendone alcuni e inserendone altri) in modo tale da cancellare quel teorema dall'insieme dei suoi teoremi, e da farvi comparire la sua negazione. Ora un'operazione di questo tipo avrebbe conseguenze potenzialmente devastanti, e in linea di principio insondabili, *sull'identità di* T, e quindi di Φ; e questo significa che Φ dovrà essere respinta globalmente, e sostituita con una tfm *sostanzialmente* diversa.

Del tutto realistico è invece il caso, già accennato più sopra e adesso meglio (ma sempre informalmente) precisato, in cui qualche assioma specifico A della tfm falsificata possa essere sostituito da un corrispondente assioma $A'(c)$ contenente certi parametri "continui" c, in modo tale che i) la tfm con $A'(c\to c_0)$ in luogo di A sia equivalente alla tfm originale, ad es. perché $A'(c\to c_0) \equiv \equiv A$, e che ii) sostituendo A con $A'(c=c^*)$, ove c^* e c_0 sono distinti ma convenientemente "prossimi" nella scala dei valori di interesse, essa diventi corretta alla luce di più precise verità empiriche. In questo caso la vecchia tfm si riduce ad un "limite" (per $c\to c_0$) della nuova. La fisica-matematica ha conosciuto situazioni clamorose del tipo appena descritto, ad es. con la transizione dall'ottica geometrica a quella ondulatoria, oppure dalla dinamica classica a quella relativistica ristretta, o ancora dalla meccanica newtoniana a quella quantistica, .. e via dicendo; ove le teorie di partenza sono limiti asintotici (per lunghezza d'onda tendente a zero in rapporto alle lunghezze tipiche, per velocità tipiche tendenti a zero in rapporto alla velocità della luce, per azioni tipiche tendenti a infinito in rapporto all'azione di Planck, ecc.) di quelle di arrivo. [30]

Continuando ad esprimerci in modo informale, diremo "solida" una tfm della quale un numero adeguatamente alto di teoremi siano stati sperimentalmente convalidati entro le precisioni osservative. Al presente, le tfm correnti sono in buona parte molto, o accettabilmente, solide. Ora è possibile immaginare situazioni in cui una tfm non sia né convalidabile né refutabile mediante esperimenti *oggi* ragionevolmente realizzabili, pur nella prospettiva che essa possa raccordarsi armoniosamente, da un punto di vista concettuale, con altre tfm solide. Sembrerebbe allora conveniente o necessario accettare quella tfm come "valida fino a prova contraria", eludendo così il dubbio di Hume-Popper in modo anche più radicale di quello correntemente praticato. L'idea di una situazione di questo tipo non è accademica, perché ad es. le presenti teorie sulla gravità quantistica sono almeno parzialmente in questa condizione. Per quanto possiamo aspettarci, una loro possibile

---

[30] Si noti che al contrario della lunghezza d'onda del primo caso, la velocità della luce e rispettivamente l'azione di Planck degli altri due sono esempi di (presunte) **costanti di natura**. Spesso le costanti di natura si fanno tendere "formalmente" a certi limiti, con ciò intendendosi che certi rapporti tipici, dei quali esse sono *uno* dei due termini, tendono ai limiti corrispondenti.

selezione su base sperimentale potrà (forse!) realizzarsi soltanto mediante procedure *molto* indirette, essendo le procedure dirette fuori della portata degli strumenti osservativi oggi disponibili o ragionevolmente immaginabili. D'altra parte un forte argomento in favore della massima apertura di principio verso i tentativi di estendere le nostre conoscenze in questa ed analoghe direzioni è il fatto che l'insieme delle attuali tfm è manifestamente *incompleto* rispetto al mondo fisico già oggi osservabile: ad esempio, e come è ben noto, il modello standard della fisica delle particelle elementari non è in grado di prevedere i valori di qualcosa come una ventina di parametri fondamentali, che devono esservi introdotti ricavandoli dall'osservazione. Questa presente stimolante situazione si accorda del resto con quel carattere dinamico ed evolutivo delle tfm che la storia delle scienze esatte costantemente ci propone come segno caratteristico della loro vitalità.

I.3) SCIENZE FORMALIZZABILI VS. SCIENZE NON FORMALIZZABILI

In pratica, la formalizzazione in T (supposto coerente) di una tfm non è mai presentata in modo *completamente* esplicito (cioè ridotta ad una successione di segni di un alfabeto elementare [31]), anche se ciò potrebbe esser fatto al prezzo di una "attività meccanica" analoga a quella del programma compilatore che traduce un testo scritto in Fortran nel linguaggio-base di un calcolatore. Questa circostanza, che rende ragione della precedente distinzione tra teorie fisiche formalizzate e semplicemente formalizzabili, è largamente giustificata da ragioni di evidenza e di economia; ma va nondimeno rimarcato che la formalizzazione esplicita di una tfm resta uno *strumento di principio* utile (se non strettamente necessario) a sondarne la "buona salute" (soundness) con il totale rigore meccanico delle procedure controllabili da un automa, oltre che l'ideale coronamento della sua definitiva messa a punto. (L'altro strumento di tale verifica essendo ovviamente la forza della ricognizione sperimentale, con buona pace dello scetticismo alla Hume-Popper.) Ma naturalmente, lo ripetiamo ancora una volta, quasi sempre non possiamo esser certi a priori, usando mezzi leali, che quella formalizzazione sia coerente; anche se probabilmente lo è.

Se da una parte il requisito della formalizzabilità di una tfm implica la sua capacità predittiva del tipo "se … allora", dall'altra non è detto che quest'ultima, o una sua "approssimazione per difetto", basti da sola a stabilire la natura teoretica di una ricerca procedente dall'osservazione; e non a caso tale capacità è condivisa, in maggiore o minor misura, da quelle

[31] Per fare un esempio significativo, affinché un automa possa maneggiare senza errori (ossia "possedere") la semplice nozione matematica di "1" (*cardinale*) occorrerebbe trasmettergliela, nel suo alfabeto elementare di pochi segni, mediante una stringa di un numero enormemente maggiore di essi (cfr. Bourbaki, "Théorie des ensembles", E III 24, ove l'alfabeto consta di 6 segni, e la stringa in questione, secondo una stima grossolana, di parecchie decine di migliaia di essi). Su questa base, sarebbe anche facile calcolare l'ordine di grandezza dell'informazione contenuta in quella nozione apparentemente elementare, espressa in bit.

scienze empiriche che non hanno ancora raggiunto il carattere di scienze esatte, o che verisimilmente non potranno mai raggiungerlo a causa della irrealizzabilità, complessità, irripetibilità, ecc., delle procedure osservative/transinduttive connesse ai loro oggetti empirici, e della difficoltà, o virtuale impossibilità, di una loro formalizzazione (**scienze** (empiriche) **non-esatte**). È fuori discussione, infatti, che una certa attitudine predittiva del tipo "se … allora" si possa attribuire ad una larga varietà di strumenti metodologici, ed al più basso livello conoscitivo a quel **principio di analogia** con cui ha inizio ogni processo di induzione; ma l'uso di questi strumenti appartiene per l'appunto al momento induttivo (o transinduttivo) sul quale anche *si conclude* il percorso delle scienze non-esatte, nella gerarchia che ha alla sua base le scienze "descrittive-elencatorie" (≡ che meramente elencano oggetti empirici e fatti intorno ad essi, avendoli univocamente identificati mediante congrue descrizioni), poi di quelle "comparative" (≡ che confrontano, sotto qualche aspetto, oggetti e fatti stabilendo così un ordine "ragionato" su di essi, generalmente parziale), di quelle "tassonomiche" o "classificatorie" (≡ che raggruppano oggetti e fatti in classi, non necessariamente disgiunte, secondo qualche criterio), .. e via dicendo.

Ben più significativo appare dunque l'insieme delle conseguenze della formalizzazione, mediante la quale una tfm *ordina* i risultati delle osservazioni in una rete logica di correlazioni interne, proponendosi di *spiegarli* in una precisa e peculiare accezione. Con queste proprietà, essa tfm funziona come un vero e proprio meccanismo di precisione; e inoltre, la sua elaborazione agisce spesso come formidabile stimolo nei confronti della invenzione matematica, al punto di spostarne il principale interesse su quello specifico terreno, prescindendo così dalla sua preminente natura di "strumento di conoscenza del mondo".

Riassumendo, la formalizzazione assegna alla generica tfm una tipica capacità di *ordinare*, *amplificare e unificare la conoscenza di* (una conveniente sezione di) *un mondo fenomenico idealizzato e presuntivamente regolare*. Con ciò, da una parte un insieme più o meno amorfo di verità empiriche, sperimentalmente convalidate entro le approssimazioni osservative, viene logicamente ordinato (ormai in un preciso senso formale) facendolo risalire, lungo le catene deduttive percorse all'incontrario, ad un piccolo numero di asserti (gli assiomi, specifici e non, convenientemente ma non univocamente individuati [32]); e dall'altra, emerge un insieme similmente

[32] Si ritiene spesso – da parte dei non-addetti ai lavori – che gli assiomi (specifici e non) siano introdotti in modo diretto e indipendente mediante le opportune procedure osservative/super-induttive (le ben note "verità evidenti" della tradizione). In realtà l'elaborazione di una tfm, e in generale di ogni sistema formale abbastanza potente, è un processo assai più tortuoso, del quale la fissazione degli assiomi (un'operazione comunque non univoca) è soltanto una parte. Insomma quella elaborazione ha nel suo insieme un carattere indiretto, non lineare, dinamico ed evolutivo: un «gioco di "avanti e indietro" noto in filosofia come "circolo ermeneutico", in cui gli assiomi sono giustificati dalla forza dei teoremi che sono in grado di giustificare, … e gli stessi teoremi, la cui verità è stata resa preventivamente evidente in modo intuitivo, sono resi inevitabili dalle dimostrazioni formali.» (G.C. Rota, in "Discrete Thoughts", Birkhäuser (1986)). Del resto il formalismo (qui inteso in senso lato) non rispecchia per nulla l'esperienza che la grande maggioranza dei matematici e dei fisici-matematici ha della pratica della propria disciplina, la quale si basa di fatto su ogni sorta di intuizioni e induzioni, spesso procedendo per tentativi. Si legga, per una critica implacabile a qualunque

ordinato di *enunciati veri nuovi*, prodotti dai meccanismi deduttivi ed eventualmente convalidabili dall'esperienza. L'enorme effetto di *sintesi* conoscitiva che da ciò consegue è ben manifesto, in quanto l'insieme di *tutti* i Teoremi della teoria, numerabilmente infinito, (a una parte dei quali corrispondono enunciati veri nella data interpretazione empirica) è generato dal piccolo numero degli assiomi e schemi di assiomi, e dall'unica regola di inferenza naturale (la regola di separazione) della teoria stessa. Infine il carattere ereditario della deduzione assicura automaticamente la *contestualità* del sistema delle verità in oggetto: esso permane cioè all'interno del contesto fissato dagli assiomi, evitando generalizzazioni insensate e/o arbitrarie.

Contestualità/specificità e potere generalizzante/predittivo sono proprietà tra loro in competizione: al crescere della prima il secondo tende a ridursi, ponendoci presto sulla strada dell'empirismo radicale, ovvero della negazione di ogni teoria. Tuttavia il fatto che le tfm esistano, e funzionino egregiamente, ci fa contare su un sufficiente spazio [33] tra quelle proprietà antagoniste. La formalizzabilità implica inoltre la natura *de-interpretabile* della relativa tfm: natura tale cioè, ad esempio, che «tutti gli enunciati di una teoria dell'elettricità valgano per ogni altro sistema di entità che si sostituiscano al posto delle nozioni di magnetismo, di carica elettrica, .. ecc., purché siano soddisfatti gli assiomi richiesti – supposti coerenti.» (D. Hilbert). [34] E infine, la stessa formalizzabilità assicura la validabilità "*automatica*" (≡ delegabile ad un automa) delle dimostrazioni dei teoremi, ciò che conferisce all'insieme degli enunciati di questi ultimi un totale grado di "certezza" (nella conveniente interpretazione) a partire da quella degli assiomi supposti coerenti, e per le date regole di inferenza. [35]

Volendo azzardare una formula conclusiva, potremmo dire che una tfm consiste nell'identificare un sistema organico e non-banale di leggi (o ipotesi) empiriche, relative ad una classe non troppo ristretta di fenomeni idealizzati, la cui essenza è nella capacità di *ordinare e amplificare sostanzialmente*, via formalizzazione, il sistema di verità nel dato contesto, empiricamente verificate o verificabili. A questo punto si sarebbe tentati di aggiungere che la natura

---

approccio dogmatico alla matematica, la tesi di dottorato di I. Lakatos "Proofs and Refutations: The Logic of Mathematical Discovery" (pubblicata a Cambridge nel 1976, ma risalente a quasi venti anni prima). Al di là di queste argomentazioni, resta il fatto che se il formalismo ben difficilmente può essere considerato come un *metodo di sviluppo* delle teorie matematiche (e tanto meno fisico-matematiche), esso è senza dubbio l'unico modo efficace di *descriverle* inequivocamente, ad impresa compiuta; senza dimenticare che considerazioni *puramente formali* hanno spesso un importante ruolo euristico nel progresso di quelle teorie.

[33] «Mentre gli assiomi più bassi non differiscono molto dalla esperienza comune …, quelli più alti e generalissimi sono meramente concettuali ed astratti … Soltanto quelli che stanno nel mezzo sono veri, solidi e vivi. » (Bacone, loc. cit., 1620).

[34] Come caposcuola della corrente formalista nella logica della prima metà del secolo XX, Hilbert fu a ragione estremamente sensibile a questa peculiare possibilità, già ben presente nel pensiero di M. Pasch nei confronti della geometria, di de-interpretare una generica tfm. La frase citata si trova in una lettera del carteggio tra Hilbert e G. Frege, e fa il paio con un'altra assai più nota espressione della stessa idea, con la quale Hilbert proponeva di sostituire i punti, le rette e i piani della sua geometria formalizzata rispettivamente con «tavoli, sedie e boccali di birra».

“scientifica” di una data attività equivalga alla possibilità di principio che essa si evolva in una tfm col progredire del suo sviluppo, anche in una prospettiva di lungo termine; [36] ma preferiamo non far qui nostra questa proposta definitoria di per sé ragionevolmente inequivoca (anche se non esente da incertezze pratiche, vedi l’esempio della nota precedente).

Possiamo tuttavia tentare qualche approfondimento in questa direzione prima di chiudere la nostra riflessione. Come dicevamo più sopra, gran parte delle nostre attività “ragionative” (e in certo senso anche psichiche) è di natura transinduttiva, cioè consiste di induzioni e possibilmente di deduzioni interne alle prototeorie formali canoniche, queste ultime per lo più a livello inconscio e nella interpretazione intuitiva. A buona ragione, si potrebbe addirittura affermare che senza questo tipo di procedure mentali la nostra stessa sopravvivenza fisica sarebbe continuamente a rischio: sono esse, infatti, che presiedono alle scelte e ai comportamenti quotidiani sulla base di stime (probabilistiche) istintive, anche se spesso inconsapevolmente raffinate. Nella misura in cui un sistema di conoscenze ottenute per via osservativa/transinduttiva viene man mano messo insieme, ordinato e organizzato, collegato e confrontato con altri analoghi sistemi, fatto oggetto di discussione e/o di insegnamento/apprendimento, e insomma istituzionalizzato, diventando così patrimonio condivisibile e trasmissibile, esso tende ad essere visto come “catalogo” o “codice” di conoscenze, poi come “sapere (in qualche modo) strutturato”, e infine come “scienza”; sebbene quest’ultima estrapolazione possa *parzialmente* confliggere con la definizione *informale* positivista che abbiamo riportato nella nota ([4]) (la diremo **definizione di scienza del tipo standard**), e possa *totalmente* confliggere con quella *formale* di scienza esatta. Ora è chiaro che mentre una data attività più o meno soddisfacente alla definizione di scienza del tipo standard ha un solo modo, precisamente stabilito, di essere anche scienza esatta (quello di essere formalizzabile), ne ha praticamente infiniti di non esserlo. In linea di principio, la vasta classe delle scienze positive (al presente) non formalizzate potrebbe essere in qualche modo ordinata in base alla misura con cui ciascuna di esse, mediante le procedure osservative/transinduttive che le sono proprie, è in grado di identificare presunte leggi non banali contestualmente generali, scoprendo così regolarità (non banali) del mondo fenomenico. Tuttavia è anche troppo evidente quanto incerti potrebbero essere i giudizi a presidio di tale ordinamento, per non dire di come la precaria gerarchia così possibilmente ottenuta risulterebbe in definitiva irrilevante agli effetti pratici. Quale utilità avrebbe infatti stabilire se al presente la clinica medica (diciamo) “sia più scienza” della vulcanologia – posto di poter dare a questa affermazione un significato obiettivo e condivisibile?

---

[35] Va da sé che qui abbiamo a che fare con due tipi di certezza, uno sintattico e l’altro semantico. In quello sintattico, si tratta della certezza meccanica della deduzione formale; in quello semantico, si tratta della certezza derivante dal teorema (™) della sezione I.1.

[36] Ad esempio, è concepibile che almeno una parte della moderna biologia molecolare sia avviata a diventare una scienza esatta; non solo, ma si cominciano a intravedere potenziali direzioni di sviluppo in tal senso, anche se in tempi per il momento sostanzialmente imprevedibili.

Vale forse la pena, avviandoci a concludere, di proporre ancora qualche commento sull'uso sempre più libero della parola "scienza" che ricorre nel linguaggio corrente contemporaneo, e che sembra attribuire una natura scientifica sui generis, rispetto ad una definizione del tipo standard, ad attività che a vario titolo con la scienza *in quella accezione* poco, e spesso nulla, sembra abbiano a che fare: alludiamo a quelle "scienze" sociali, politiche, storiche, teologiche, occulte, della comunicazione, del comportamento, della formazione, .. e via seguitando, che così insistentemente sono nominate nel parlare quotidiano. Al di là dell'innocua questione nominalistica, che al più testimonia la drammatica mancanza di comunicazione tra mondi culturali separati anche nel linguaggio, *è bene riflettere sulle distinzioni di sostanza e trarne le dovute conclusioni*. Sistemi di conoscenze come quelli sopra ricordati – non formalizzabili salvo che in alcuni rari casi e sotto loro aspetti particolari [37] – sono tutti assai lontani, nei metodi e nei fini, da quello scientifico-esatto; e ciò, non soltanto perché la loro elaborazione non soddisfa nemmeno le richieste minimali previste da una definizione di scienza del tipo standard, o le soddisfa assai imperfettamente. Infatti, le procedure osservative/transinduttive delle quali l'elaborazione di un eventuale modello teorico consiste (quando esistono), di norma *non sono mirate* ad individuare presunte leggi non banali e contestualmente generali (non raggiungendo quindi nemmeno il livello dell'induzione in senso stretto); e in conclusione, esse *non possono condurre, e di fatto non conducono, alla scoperta di regolarità non banali del mondo*. Se si accettano queste ragionevoli argomentazioni, ben difficilmente i sistemi di conoscenze della precedente lista (e simili) possono considerarsi come "scienze" (quanto meno nel senso della definizione standard); sebbene l'uso corrente le denomini con la stessa parola. Dal presente punto di vista, siamo cioè dinnanzi ad un tipico caso di "abuso di linguaggio" che si realizza e si accetta correntemente all'interno di un *gergo*, talvolta radicato nella tradizione e tal'altra assai più recentemente introdotto. [38]

Ci limitiamo qui ad esaminare in qualche dettaglio una situazione esemplare, caratterizzata dall'uso della cosiddetta (dai logici) "logica del detective": quella della ricerca storica. Ovviamente l'esercizio di questa logica "interpolativa" (un aspetto di quella induttiva) non è estraneo al processo scientifico-esatto – del quale è anzi quasi sempre un componente importante –, ma è ben lontano dal costituire di per sé un atto di quel tipo, o anche una sua premessa. Infatti ogni caso

---

[37] Come ad esempio in quello di certi meccanismi competitivi presenti nei o tra i gruppi umani, modellizzabili e matematizzabili all'interno della Teoria dei Giochi, e atti a rivelare possibili "sezioni esatte" (ideali) delle cosiddette scienze sociali, politiche e perfino storiche.

[38] È invece ben naturale un'interpretazione maliziosa delle ragioni che spingono le comunità interessate a rivendicare il preteso carattere scientifico delle loro attività o discipline: per comune ma raramente esplicito riconoscimento, la scienza – quella vera – ha ormai infatti un impatto formidabile sulla società contemporanea, sia sul piano culturale che su quello delle sue ricadute socio-economiche, e conferisce in modo quasi automatico autorevolezza a chi la produce, o più semplicemente la sostiene o la utilizza. Ben di rado questa autorevolezza si traduce in autorità o in vero e proprio potere; ma per ragioni in parte inspiegabili, ugualmente essa richiama e soggioga una parte importante della pubblica attenzione.

poliziesco sta a sé, ed il fine del detective è di regola quello di risolvere quel caso specifico, volta per volta, e non certo quello di *indurre* leggi probabilistiche generali da una statistica di casi polizieschi per *dedurne* un sistema sostanzialmente amplificato di altre leggi probabilistiche generali. Se poi per mera ipotesi ciò si verificasse con le necessarie garanzie di universalità contestuale, rigore definitorio/deduttivo e non-banalità, allora a tutti gli effetti quel detective starebbe tentando di elaborare una "teoria probabilistica formalizzabile del caso poliziesco"; e quest'ultima attività, quand'anche seriamente concepibile, sarebbe cosa ben diversa dalla prevedibile sua finalità originaria (vale a dire, il nostro detective sarebbe sul punto di cambiare mestiere). Con ciò non si nega che il detective possa farsi una sua "esperienza generale (non banale [39]) sulla criminalità", ma soltanto che questa esperienza sia formalizzabile come lo è nelle scienze esatte, e che quindi possa offrirgli certezze predittive di qualche valore. La metafora, applicata all'attività tipica dello storico, è trasparente. [40]

Ci piace chiudere con un noto aforisma di E. Rutherford, che recita: «una scienza o è fisica [41] o è filatelia». Al di là del suo gusto per le battute, in questa occasione il celebre Nobel inventore del modello atomico (1911) non scherzava, e anzi si dimostrò di manica larga. Infatti "filatelia" è una ovvia metafora per "classificazione"; ma la maggior parte delle soprannominate cosiddette "scienze" non raggiunge nemmeno il livello della tassonomia.

[39] Un asserto banale sulla criminalità è ad esempio: "La criminalità aumenta in condizioni di disordine sociale" (o piuttosto è probabile che aumenti). È facile immaginare un gran numero di amenità di questo tipo.

[40] La pretesa scientificità della Storia è difesa (in verità assai debolmente nel metodo e nelle conclusioni, queste ultime quasi sempre di sconcertante ovvietà) da certe correnti filosofiche contemporanee.

[41] È chiaro che "fisica" è qui da intendere nel senso di "fisica formalizzabile". In apparenza in modo indipendente, la sentenza di Rutherford è stata riproposta quasi alla lettera, in tempi recenti e forse ormai un po' sopra le righe, da G.C. Rota (nei già citati "Discrete Thoughts"), prendendo a termine di confronto la biologia: «Purtroppo il Galileo della biologia non è probabilmente ancora nato; e nell'attesa, le metodiche biologiche poco si discostano da quelle di un collezionista di francobolli.» Ci risulta poi che anche il fisico L. Alvarez (morto nel 1988) abbia ripreso l'immagine in una polemica in cui manifestava disprezzo verso i paleontologi e i loro contributi alla scienza: «Davvero (i paleontologi) non sono buoni scienziati. Più che altro sono dei collezionisti di francobolli.», ebbe a scrivere in un duro articolo sul New York Times (citato in B. Bryston, "A Short History of Nearly Everything", 2003). Insomma l'unica

BIBLIOGRAFIA GENERALE

BOURBAKI, N.: "Theorie des Ensembles", Hermann, 1970

CARNAP, R.: "Philosophical Foundations of Physics", Basic Books, 1966; trad. it. "I fondamenti filosofici della fisica", Il Saggiatore, 1971

CASARI, E.: "Introduzione alla logica", UTET 1997

CHANG, C.C., KEISLER, H.J.: "Model Theory", North Holland 1973

FRAENKEL, A., BAR-HILLEL, Y.: "Foundations of Set Theory", North Holland 1973

FRANK, P.: "Philosophy of Science: The Link between Philosophy and Science", Prentice Hall 1957

KLEENE, S.: "Introduction to Metamathematics", North Holland 1952

KLINE, M.: "Mathematics: The Loss of Certainty", Oxford Un. Press 1980

NAGEL, E.: "La struttura della scienza" (trad. ital.) Feltrinelli 1968, ed. origin. 1961

POINCARÉ, H.: "Science et Méthode", Flammarion, 1908 ; "La Science et l'Hypotèse", Flammarion, 1920; entrambi in trad. ital. in "Opere epistemologiche", Piovan 1989

ROBINSON, A.: "Introduction to Model Theory and to the Metamathematics of Algebra", North Holland 2nd ed. 1965

RUSSELL, B.: "Our Knowledge of the External World" (various printings, 1926), trad. ital. "La conoscenza del mondo esterno", Longanesi 1966

SCHÜTTE, K.: "Proof Theory", Springer 1977

SUPPES, P.: "Studies on the Methodology and Foundations of Sciences", Reidel 1969

---

cosa che sembra chiara, in questa querelle, è che l'attività dei poveri filatelisti non sia tenuta in gran conto da fisici e matematici …

II – INTRODUZIONE ALLA TEORIA ANALITICA DELL'APPROSSIMAZIONE

II.1) GENERALITÀ

Più o meno direttamente, abbiamo fin qui delineato due anime della scienza esatta teorica: quella di chi disegna strategie interpretative, o modelli, (di sezioni) del mondo fenomenico, e quella di chi quei modelli formalizza, inventandone e soprattutto studiandone la matematica. Una parte cospicua della seconda di queste attività è andata via via specializzandosi ed acquistando crescente importanza e autonomia a partire dal XVII-XVIII secolo. Essa consiste nel (cercare di) risolvere le equazioni incorporate nei modelli fenomenici (o in generale equazioni qualsiasi); e con questo, emerge una terza e assai peculiare anima della scienza esatta teorica (nonché della stessa matematica), alla quale non abbiamo finora riservato l'attenzione che merita.

Nella sezione I.1 del precedente articolo (I) abbiamo dato la più generale definizione di "equazione" già possibile in un sistema formale (sf) molto debole (una teoria logica-quantificata egualitaria). Una definizione più naturale, che rientra nel caso precedente, è possibile in teoria degli insiemi, ed è la seguente. Si consideri l'applicazione f: $D \to C$, dove D e C sono insiemi arbitrari, e sia F il suo grafo (funzionale). Per ogni sottoinsieme c di C, l'immagine $F^{-1}[c]$ di c attraverso F^{-1} (grafo reciproco di F) è un sottoinsieme di D, possibilmente vuoto, che si dice "immagine reciproca di c attraverso F". Per un dato $y \in C$, "risolvere l'equazione

(1) $f(x) = y$"

significa "identificare l'immagine reciproca di $\{y\}$ (singoletto di y) attraverso F", o almeno una sua parte. Ovviamente gli elementi di $F^{-1}[\{y\}] \subset D$ sono le **soluzioni** della (1). Dovrebbe esser chiaro che soltanto in casi molto rari è possibile risolvere la (1) mediante un numero finito di passaggi logici: presupponendo che D sia uno *spazio metrico completo*, e quindi che sia possibile introdurvi il concetto di *soluzione approssimata*, la "via del re" verso questo obiettivo consiste nell'ottenere soluzioni approssimate ma di accuratezza arbitraria della (1), fino a quando l'errore stimato sia considerato accettabile.

Con questo articolo (II), proponiamo al lettore una breve incursione informativa nel mondo della **teoria analitica dell'approssimazione**, al cui interno la risoluzione (di accuratezza finita, ma arbitrariamente alta) di equazioni occupa una posizione di rilievo assoluto. Considerate le attuali dimensioni di quel mondo, l'impresa è a dir poco temeraria. L'impressione fondamentale che si ricava da un esame di largo orizzonte della "scienza equazionale" è infatti al più quella di un insieme di strategie mirate a risolvere una certa equazione muovendo da particolari altre equazioni ad essa convenientemente “prossime”, e di essa più facilmente risolvibili con la precisione desiderata. Ovviamente ci si aspetta che le soluzioni di queste equazioni “approssimanti” siano a loro volta prossime a quella dell’equazione di partenza. È poi evidente che questa descrizione richiede molte e non banali precisazioni tecniche per acquistare lo status di una definizione matematica.

Di fatto, qualunque approccio didattico moderno alla teoria della risoluzione di equazioni non può che partire dalla analisi funzionale cosiddetta “non lineare” − o più propriamente dall’analisi funzionale senza altri attributi. [42] Poiché siamo principalmente interessati alle equazioni che traggono origine da teorie fisiche, sarà utile fare un passo indietro per dare al problema un conveniente inquadramento generale. Torniamo dunque alla (5) della sezione I.1, cioè alla versione insiemistico-empirica di una generica legge fisica. Questa sarà qui riscritta come relazione *assertiva* su x

(2) “x appartiene a $\mathcal{D}$, e $\mathcal{P}(x)$”,

dove $\mathcal{P}$ è un predicato e x è un individuo a priori appartenente a un dato universo $\mathcal{D}$ di osservabili. Ovviamente “x appartiene a $\mathcal{D}$, e $\mathcal{P}(x)$” è collettivizzante in x, cioè individua unicamente un insieme incluso in $\mathcal{D}$ che denotiamo con $\mathrm{Ext}(\mathcal{P}|_{\mathcal{D}})$, l’estensione di $\mathcal{P}$ in $\mathcal{D}$. La (1) *equivale* così all’asserto

(3) “x appartiene a $\mathrm{Ext}(\mathcal{P}|_{\mathcal{D}})$”.

Avendo identificato $\mathrm{Ext}(\mathcal{P}|_{\mathcal{D}})$, il grado di convalida empirica della (2) ci dà una commisurata fiducia di poter “prevedere” che il generico oggetto osservabile x appartenente a $\mathcal{D}$ renda vero il predicato $\mathcal{P}(x)$ semplicemente verificando che esso appartiene a $\mathrm{Ext}(\mathcal{P}|_{\mathcal{D}})$, quindi senza aver osservato la “verità fattuale” della (2), e cioè di $\mathcal{P}(x)$. Ma come sappiamo (vedi I.1), in pratica la possibilità di identificare $\mathrm{Ext}(\mathcal{P}|_{\mathcal{D}})$ passa per la *formalizzazione* della (2) in un conveniente sistema formale (sf) T includente la teoria degli insiemi. Con questo ci si sposta sulla immagine di tipo (I.1, 5ter) della (2) in T, cioè sull’asserto

(4) “$(x\in D\wedge P(x))$ ha valore di verità$_M$* uguale a 0”,

dove $M \equiv \langle \mathcal{E},I\rangle$, $\mathcal{E}$ è la struttura insiemistica empirica entro la quale si colloca la (2), I: $[\mathcal{E}] \to_{ini} [T]$ è l’iniezione che porta da $[\mathcal{E}]$ in [T] (cfr. I.1, 1), e P, x, e D sono nell’ordine le immagini (o “interpretazioni”) in T di $\mathcal{P}$, x e $\mathcal{D}$ sotto I. Il teorema (™) di cui alla sezione I.1 afferma allora che la (4) è soddisfatta se

---

[42] L’attributo “non lineare” è entrato nell’uso da molto tempo, ed è stato recentemente “santificato” dal titolo del monumentale trattato di E. Zeidler (“Nonlinear Functional Analysis and its Applications”, 1985). Secondo questa consuetudine, l’oggetto matematico che potrebbe essere lineare, *ma la cui linearità non è esplicitamente assunta*, viene detto “non lineare”, e così impropriamente presentato come complemento della sua controparte lineare. Applicato all’analisi funzionale, quest’uso (che rifiutiamo) ha certamente contribuito a porre questa disciplina in una luce riduttiva ed iniziatica agli occhi del non specialista, il quale raramente ne percepisce la natura centrale e fondante.

(4bis) “$(x \in D \wedge P(x))$ è deducibile in T”;

ovvero se l’equivalente relazione di appartenenza

(4ter) “$x \in \mathrm{Ext}(P|_D)$”,

dove $\mathrm{Ext}(P|_D)$ è l’estensione di P in D (inclusa in D), è un teorema di T. A questo punto il fisico teorico correttamente afferma di avere *formalizzato* in T la legge fisica (2): vale a dire, egli lavora ormai soltanto sulla immagine formalizzata $x \in D \wedge P(x)$ di quella legge, immagine della quale cercherà di determinare l’estensione $\mathrm{Ext}(P|_D)$. Adottando la precedente definizione (1) di equazione, il problema diventa dunque quello di identificare l’estensione (inclusa in D) della relazione in x $(x \in D) \wedge (f(x)=y)$, per y dato in C, supposta deducibile in T. È precisamente in connessione con questo problema che nasce quella particolare specie di matematico, o forse di matematico applicato, che potremmo qui nominare come “risolvitore di equazioni” (in particolare, di equazioni fenomenologiche).

Prima dell’invenzione del calcolo differenziale, tanto il dominio D delle potenziali soluzioni della (1) quanto l’insieme di arrivo C erano insiemi *numerici*, cioè erano parti (di solito aperte) della retta reale o del piano complesso, o di loro potenze cartesiane − anche se all’epoca non si aveva una precisa cognizione di tali oggetti −, e molte delle equazioni che si consideravano erano di tipo algebrico; cioè la f della (1) era un polinomio (possibilmente a coefficienti complessi) in una indeterminata reale o complessa, che veniva uguagliato a 0 (zero della retta o del piano), o un sistema di tali polinomi in altrettante indeterminate. Lo sfortunato genio di E. Galois riscoprì per suo conto (già P. Ruffini [43] e N. Abel [44] erano giunti ad analoghe conclusioni) che le sole equazioni algebriche in una incognita passibili di soluzione generale “per radicali”, o “esatta”, erano quelle di grado inferiore al quinto. [45] Per quanto qui ci interessa, va subito detto che se da una parte le rimanenti equazioni algebriche (cioè quelle di grado ≥ 5) non hanno mai avuto un ruolo importante in fisica matematica, dall’altra la possibilità di risolverle “approssimativamente” (in un senso da precisare) con metodi generali − soprattutto dopo il grande teorema gaussiano che assicurava l’esistenza di esattamente n soluzioni della equazione algebrica di grado n arbitrario (in campo

[43] P. Ruffini: “Teoria generale delle equazioni, in cui si dimostra impossibile la soluzione algebrica delle equazioni generali di grado superiore al quarto”, Bologna (1799). Nel linguaggio della moderna teoria dei gruppi di sostituzioni, il risultato di Ruffini si esprime dicendo che il gruppo delle sostituzioni su 5 lettere non possiede sottogruppi di indice 8, 4 e 3.

[44] N.H. Abel, in Journ. für Math., I 65-84 (1826). Abel dimostrò che le radici di una equazione risolvibile “per radicali” si possono porre in forma tale che ciascun radicale che vi compare sia esprimibile come funzione razionale delle radici dell’equazione. Usando questo fatto, egli riescì a dimostrare l’impossibilità di risolvere per radicali l’equazione generale di grado superiore al 4°. La dimostrazione di Abel fu poi sostanzialmente semplificata da Kronecker oltre cinquanta anni più tardi.

[45] Galois inquadrò le sue riflessioni in quella che oggi chiamiamo “Teoria dei Gruppi di Sostituzioni”, e che egli inventò quasi dal nulla. Il concetto di gruppo era già apparso a latere di problemi diversi, ma con Galois entrò a pieno diritto tra gli oggetti fondanti della matematica, aprendo la strada ad una concezione più astratta ed unitaria dell’algebra. Questo è certamente il più grande e inestimabile contributo di Galois alla matematica.

complesso, e non necessariamente distinte) – ha certamente stimolato la messa a punto di procedure di soluzione approssimata di equazioni *generiche*, algebriche e non. Va tenuto presente, in proposito, che già all'inizio del XVIII secolo erano disponibili algoritmi per il calcolo, con precisione arbitraria, di numerose funzioni reali definite in un intervallo reale (e talvolta complesse e definite in un dominio del piano complesso) e quindi di loro combinazioni funzionali ($\circ$), come ad es. sinlog, non troppo complicate. Detta f = f(x) una funzione ottenuta in tal modo, diventava dunque assai naturale proporsi di calcolare le soluzioni (a priori appartenenti al dominio $D \equiv D(f)$) di una equazione del tipo (1), per y dato reale o complesso a seconda del caso. In effetti, procedure generali finalizzate a questo scopo cominciavano ad essere proposte e messe a punto proprio allora: basti ricordare il metodo iterativo di Newton detto "delle tangenti", che si riferisce ad una f reale definita su un intervallo reale, quello di D. Bernoulli per la risoluzione iterativa di equazioni algebriche di grado arbitrario, .., ecc.

Ma in realtà è soltanto con l'invenzione di quelle particolari equazioni *funzionali* che sono le equazioni differenziali che la teoria dell'approssimazione entra nella sua età adulta. Il dominio D delle potenziali soluzioni diventa con ciò un *insieme di funzioni*: dapprima di funzioni tra spazi numerici (ad esempio delle funzioni reali definite su (0,1) e congruentemente regolari), e più tardi di funzioni di oggetti arbitrari, ma comunque appartenenti a uno spazio *lineare*, e con valori anch'esse in uno spazio lineare. Nel linguaggio matematico moderno, l'equazione stessa assume la forma:

(5) $f(x) = O_Y$,

dove f è un'applicazione di $D = D(f)$ (dominio di f) $\subset X$ in Y (quindi con $x \in D$ e $f(x) \in Y$), e X, Y sono spazi lineari (entrambi su **R** o entrambi su **C**), di solito normati e completi (spazi di Banach, o B-spazi), x è l'incognita, e O_Y è lo zero di Y. (La specializzazione del 2° membro della (5) in O_Y è chiaramente inessenziale, e non ne limita la generalità.) Per fare qualche esempio, la (5) può essere un'equazione differenziale ordinaria o alle derivate parziali, un'equazione integrale o integro-differenziale, un sistema di equazioni dell'uno o dell'altro tipo, .. e via dicendo.

All'inizio del secolo XVIII, la natura delle equazioni subisce dunque una silenziosa ma radicale evoluzione: accanto alle tradizionali equazioni con incognite numeriche *ne compaiono altre aventi per incognite delle funzioni*. Alla radice del fenomeno ci sono le equazioni differenziali ordinarie, e in primo luogo l'equazione fondamentale della dinamica estesa a sistemi di punti materiali in mutua interazione. In particolare con la teoria della gravitazione universale (Newton, 1687), per la prima volta nel suo percorso conoscitivo verso il mondo fenomenico il fisico matematico dispone di un modello formalizzato di una proprietà fondamentale di quel mondo (e della cui validità sembra irragionevole dubitare), ma deve al contempo constatare *la sua sostanziale incapacità* di estrarne quanto in effetti esso contiene in termini di potere predittivo. Il riferimento al problema degli $N > 2$ corpi in interazione gravitazionale è ovvio. Come sappiamo, il modello si formalizza in un sistema di 2N equazioni differenziali ordinarie del 1° ordine in forma normale, sotto le convenienti 2N condizioni accessorie. Il problema consiste quindi nel costruirne la soluzione generale "esatta" (cioè esprimibile in termini di funzioni note e/o di loro primitive); ma l'impresa si rivela praticamente impossibile (per $N > 2$) con i mezzi dell'analisi del tempo, ed anche, come diverrà chiaro più tardi, con mezzi molto più sofisticati.

Vale la pena di sottolineare una volta di più la novità della situazione: con la teoria della gravitazione, il fisico-matematico del Settecento dispone di un modello formalizzato di estremo interesse fisico e filosofico, ma non è in grado di coglierne che una minima parte delle implicazioni "esatte" *a causa della sua limitata capacità di calcolo*. Con pochissime eccezioni, egli deve quindi sistematicamente ricorrere a metodi approssimati per la determinazione delle soluzioni dei problemi dinamico-celesti. Col tempo, questo tipo di difficoltà si ripresenterà in un numero sempre maggiore di occasioni diverse, fino a manifestarsi come "regola naturale del gioco". D'altra parte la soluzione di equazioni è a questo punto un obiettivo irrinunciabile per il matematico applicato: perché un conto è sapere che certi individui $x \in D$ soddisfano un certo predicato equazionale P, e un conto è conoscere l'insieme $\mathrm{Ext}(P|_D)$ al quale tutti e soli quegli individui appartengono. Detto in termini crudi, un conto è *scrivere* un'equazione, e un conto è *scriverla e risolverla*.

Con poche eccezioni precorritrici (pensiamo ad esempio all'opera di L. Pisano, o al calcolo di irrazionali come π), è dunque essenzialmente sotto la spinta dei problemi dinamico-celesti non altrimenti dominabili che prende avvio quell'immenso corpo della moderna matematica applicata noto come Teoria Analitica dell'Approssimazione. Nelle analoghe istanze che si produrranno in seguito, nell'ambito della fisica matematica, il momento modellizzante-fondativo andrà sempre più distinguendosi, se non separandosi, da quello matematico-calcolatorio. Per fare un esempio relativamente recente, a partire dal terzo decennio dello scorso secolo un'altra spinta in questa direzione venne dalla meccanica quantistica; e non pochi fisici teorici furono forzati dalle circostanze a trasformarsi in "risolvitori" delle equazioni di quella meccanica, in particolare mediante i cosiddetti "metodi perturbativi", fondati su sviluppi in serie formali, o asintotiche, di potenze di uno o più piccoli parametri. [46]

In effetti esistono interi importanti capitoli della fisica teorica, diversi dalla dinamica celeste e più anziani della meccanica quantistica, i cui problemi sono (o dovrebbero essere) vere e proprie "glorificazioni" del calcolo approssimato e dell'analisi asintotica. Basti l'esempio della dinamica dei fluidi barotropici classici (ideali o dissipativi), o della magnetofluidodinamica (dinamica dei fluidi elettricamente conduttori accoppiata alle equazioni di Maxwell di bassa frequenza); o per

---

[46] Per la precisione, i metodi perturbativi esistevano già nel tardo Settecento (Lagrange), ed erano addirittura applicati alla soluzione approssimata di problemi generalmente *non lineari*, quali *non* sono quelli quanto-meccanici.

spingerci più avanti, quello della fisica teorica dei plasmi, sia astrofisici che di laboratorio. Secondo una poco sensata locuzione gergale, teorie fisiche di questo tipo sono usualmente classificate, dal punto di vista matematico, come "altamente" (?) non lineari; e per l'appunto, l'approccio alla soluzione approssimata delle loro equazioni è tipicamente inteso ad aggirarne la non linearità, trasformandole in convenienti gerarchie (di norma troncate al più basso ordine significativo, l'"ordine uno") di equazioni lineari salvo che all'ordine più basso, o "ordine zero". Riprendendo il caso della fisica teorica dei plasmi, il problema matematico consiste nel risolvere un sistema di equazioni integro-differenziali parziali non lineari (sistema di equazioni bilineari di Boltzmann, tante quante le *specie* di particelle considerate), ancora accoppiato con il sistema di Maxwell di bassa frequenza. I progressi su questo fronte sono legati a drastiche ipotesi semplificatrici e ad un tipo di analisi asintotica più o meno euristica; o infine, in tempi recenti, all'uso massiccio del calcolatore. Un contributo matematico importante è dovuto a Hilbert, che all'inizio del secolo scorso propose un metodo perturbativo fondato sullo sviluppo delle funzioni di distribuzione in serie di potenze di piccolo parametro − nel cosiddetto "limite delle collisioni dominanti" − proprio allo scopo di sostituire il sistema bilineare di Boltzmann con una gerarchia di sistemi lineari salvo che all'ordine zero (metodo di Hilbert, ulteriormente sviluppato da Enskog [47]).

Per concludere: non è irragionevole sostenere che la moderna analisi applicata si divida ormai in due grandi corpi, il primo coincidente con la teoria dell'approssimazione, e il secondo − non spropositatamente più ampio − che ne comprende tutto il resto. Lungi dall'essere una battuta ad effetto, questa affermazione descrive una realtà carica di significati e di assai concrete conseguenze.

II.2) ALCUNE APPLICAZIONI

È opportuno partire con l'illustrazione di qualche esempio. Cominceremo col descrivere un tipo di calcolo approssimato la cui portata va talvolta in effetti al di là di quanto si potrebbe stimare a prima vista. Si supponga di voler risolvere un'equazione del tipo (II.1, 1), che numereremo come la (1) di questa sezione, dove $f: D \to Y$ è una funzione di dominio $D \subset X$ in Y, X e Y essendo due spazi lineari, diciamo su $\mathbf{R}$, $x \in D$ è l'incognita, e $y \in Y$ è dato. Si supponga poi che f possa scriversi nella forma $f_0 + f_1$, dove f_0 e f_1 sono anch'esse funzioni da D in Y, e che l'equazione

(2) $\quad f_0(x) = y$

sia risolvibile $\forall y \in f_0(D) \subset Y$ (cioè che f_0 abbia un inverso f_0^{-1} noto). È allora assai naturale, detta

[47] D. Hilbert, Gött. Nachr. 355 (1910), Math. Annalen, **72**, 562 (1912), "Grundzüge einer allgemeine Theorie der linearen Integralgleichungen", Teubner 1912; D. Enskog, "The Kinetic Theory of Phenomena in Fairly Rare Gases",

(2′) $x_0 =: f_0^{-1}(y)$

la soluzione della (2), e se $f_0(x_0) \neq y - f_1(x_0)$, scrivere l'equazione

(3_0) $f_0(x_1) + f_1(x_0) = y$,

in una nuova incognita x_1 determinabile come $x_1 = f_0^{-1}(y-f_1(x_0))$ se $y - f_1(x) \in f_0(D)$. Se poi $f_0(x_0) = y - f_1(x_0)$, x_0 è evidentemente una soluzione della (1). Proseguendo in modo analogo, si daranno tre casi: o la gerarchia di equazioni (per $n \geq 0$)

(3_n) $f_0(x_{n+1}) + f_1(x_n) = y$,

se $y - f_1(x_n) \in f_0(X)$ equivalenti una ad una alle

($3_n'$) $x_{n+1} = f_0^{-1}(y-f_1(x_n))$,

genera una successione $x_0, x_1, .., x_n, \ldots$ per la quale esiste un valore $\underline{n}$ di n tale che $x_{\underline{n}+1} = x_{\underline{n}}$, oppure un tale valore $\underline{n}$ non esiste; oppure la (3_n) non è risolvibile rispetto a x_{n+1} (perché $y - f_1(x_n) \notin f_0(X)$). Nel primo caso la soluzione è $x_{\underline{n}}$; nel secondo nasce il problema della possibile "convergenza" (per $n \to \infty$ e per il dato y) della successione generata dalle ($3_n'$); e infine nel terzo ci si deve fermare all'approssimazione x_n (che non è detto sia migliore di x_0). In pratica, il problema ha senso soltanto se X e Y sono spazi di Banach (B-spazi), o almeno spazi metrici completi. Se la successione x_0, x_1, .., .. è univocamente costruibile, e converge [48] per il dato y, detto x il suo limite, e *supposte le* f_0, f_1 *continue*, la (3_n) diventa

(3_∞) $\lim_{n\to\infty} [f_0(x_{n+1}) + f_1(x_n)] = f_0(x) + f_1(x) \equiv f(x) = y$;

cioè x è soluzione del problema.

Si noti ancora che non è necessario, ma soltanto ragionevole, partire con la $x_0 = f_0^{-1}(y)$ nella costruzione della successione $x_0, .., x_n$, .. Qualunque altra $x_0' \in X$ (che genererà la successione x_0', .., x_n', ..) è a priori accettabile; anzi, può avvenire che la prima successione non converga e la seconda converga, o che valga qualsiasi altra combinazione dei due casi. Si vede cioè che il successo finale della procedura – la convergenza, per il dato y, della successione generata – dipende dalla scelta della decomposizione di f in $f_0 + f_1$ (continue, e la prima con range abbastanza ampio e invertibile) e della x_0 "di primo tentativo", e da niente altro.

Se pensiamo y come fisso, la precedente descrizione del problema è ridondante, nel senso che $-y$ può incorporarsi in f, passando con questo all'equazione $f(x) = O_Y$. Allora anche l'equazione $f_0(x) + f_1(x) = 0$ (scriviamo ormai O_Y come 0) è ridondante se f_0 è supposta invertibile e con range abbastanza ampio, perché essa può scriversi come $x = -f_0^{-1}(f_1(x)) \equiv g(x)$. Per ragioni ovvie, un'equazione di questo tipo, con $g: D(g) \subset X \to X$ (ove $D(g) \equiv$ dominio di g, e X è un B-spazio), si

Dissertation, Uppsala (1917). Si veda anche il classico trattato di S. Chapman e T. Cowling, "The Mathematical Theory of Non-uniform Gases", Cambridge Un. Pr. 1953.

dice (equazione) **di punto fisso**, ed è alla base di molti importanti algoritmi di calcolo approssimato. Un'equazione di punto fisso suggerisce immediatamente la catena di sostituzioni successive ($^+$) $x_{n+1} = g(x_n)$, caso particolare della ($3_n'$). In generale, vi è più di un modo per trasformare un'equazione del tipo $f(x) = 0$ in una equazione di punto fisso, e la convergenza della risultante successione ($^+$) può essere influenzata dalla scelta operata.

Partendo da queste considerazioni, ci proponiamo ora di risolvere l'equazione differenziale ordinaria del 1° ordine, in forma normale e generalmente non autonoma,

(6) $d_t x(t) = F(x(t),t)$,

sotto la condizione iniziale

(6′) $x(0) = 0$,

dove F è una funzione continua di $\mathbf{R}^2$ in $\mathbf{R}$ (con le norme standard). Le (6,6′) si trasformano subito nella equivalente equazione integrale (di punto fisso)

(7) $x(t) - \int_{t=0}^{t} F(x(t'),t'))dt' = 0$,

alla quale possiamo applicare l'algoritmo ($2',3_n'$). Detto cioè f(x) il 1° membro della (7), possiamo identificare $f_0(x)$ della precedente procedura con l'identità su **R**, e $f_1(x)$ con $-\int_{t=0}^{t} F(x(t'),t'))dt'$, mentre la y è nulla. La $f_0(x_0) = 0$ dà allora $x_0 = 0$, la ($3_0'$) diventa $x_1(t) = \int_{t=0}^{t} F(0,t')dt'$, e la ($3_n'$) diventa $x_{n+1} = \int_{t=0}^{t} F(x_n(t'),t')dt'$. Il calcolo della successione $x_0 = 0, x_1, .., x_n, ..$ è così *ricondotto a quadrature*; se la successione converge, il suo limite è soluzione della (7), e quindi delle equivalenti (6,6′). Questo metodo di soluzione di una equazione differenziale del tipo (6) sotto la (6′) (o con una modifica banale sotto una generica $x(0) = a$), è detto (metodo di soluzione) **per successive approssimazioni**. La teoria generale è dovuta a É. Picard, [49] per cui esso è anche noto come **metodo** (o **algoritmo**) di **Picard**, in concorrenza con quello alle differenze di (Eulero)-Cauchy-Lipschitz.

Tornando alla (1), l'algoritmo ($2',3_n'$) può essere immediatamente visto come procedura per la sua soluzione mediante **sviluppo in serie formale di potenze di piccolo parametro**, o **metodo perturbativo**, se f_0 e f_1 sono anche *lineari*, f_0 restando comunque invertibile (la richiesta di linearità essendo sufficiente, ma non necessaria). In luogo della (1), consideriamo la

(4) $f_0x + \varepsilon f_1x = y$,

---

[48] Cioè se per dato $\varepsilon > 0$ arbitrario, esiste un $N = N_\varepsilon$ tale che per ogni n e m maggiori di N_ε,, $||x_n - x_m|| < \varepsilon$ (criterio di Cauchy). In particolare la convergenza della successione implica che $||x_{n+1}-x_n|| \to 0$ per $n \to \infty$.

[49] É. Picard, Journ. Math. Pures et Appl., 6, 145-210 (1890). La teoria di Picard è riassunta nel cosiddetto **teorema di Picard-Lindelöf**. Più o meno la stessa procedura era stata tuttavia applicata molto tempo prima, su basi e con obbiettivi meno generali, da J. Liouville.

dove $\varepsilon \in [0,1]$ (quindi per $\varepsilon = 1$ la (4) coincide con la (1) con funzioni f_0, f_1 lineari). Sia $x(\varepsilon)$ una soluzione della (4) che rappresentiamo tentativamente come serie di potenze di ε, $x(\varepsilon) = \sum_{i=0} \varepsilon^i x_i^*$. Evidentemente, $x_0^* = x(\varepsilon{=}0) = f_0^{-1}(y) = x_0$. Sostituendo tale serie nella (4), abbiamo $\sum_{i=0} \varepsilon^i f_0 x_i^* + {} + \sum_{i=0} \varepsilon^{i+1} f_1 x_i^* = y \ \forall \varepsilon \in [0,1]$. In forza del principio di identità polinomiale, oltre alla $x_0^* = x_0$ otteniamo così la gerarchia di uguaglianze

(5_n) $\quad f_0 x_{n+1}^* + f_1 x_n^* = 0$,

per $n \geq 0$. Queste coincidono con le (3_n) se identifichiamo x_n con $\sum_{i=0}^n x_i^*$, cioè con il valore della somma parziale $\sum_{i=0}^n \varepsilon^i x_i^*$ per $\varepsilon = 1$. Per $n = 0$, abbiamo infatti $0 = f_0 x_1^* + f_1 x_0^* = f_0(x_1 - x_0) + f_1 x_0 = {} = f_0 x_1 + f_1 x_0 - y$, che è la ($3_0$) con f_0 e f_1 lineari; e similmente, $\forall n > 0$, $f_0 x_{n+1} + f_1 x_n = f_0 x_n + f_1 x_{n-1} = \ldots = f_0 x_1 + f_1 x_0 = y$, che sono le ($3_n$), qed. Ciò spiega perché l'algoritmo ($2'$,$3_n'$) – valido per f_0, f_1 continue ma non necessariamente lineari, e f_0 biiettiva su $f_0(D)$ – è talvolta presentato come "metodo perturbativo". Se per il dato y la serie $\sum_{i=0} \varepsilon^i x_i^*$ converge per $\varepsilon \to 1-$, il suo valore è soluzione della (1). Al di fuori del caso lineare (4), i metodi perturbativi richiedono in generale l'uso di strumenti matematici più avanzati (il calcolo differenziale *astratto*, cioè in B-spazi).

All'algoritmo di Picard si riconduce il cosiddetto **metodo di variazione delle costanti**, che è stato presto usato per la soluzione ricorsiva dell'equazione di Hamilton-Jacobi (H.J.) Consideriamo una hamiltoniana n-dim $H(t,q,p)$ e le associate equazioni canoniche

(8_1) $\quad d_t q = \partial H/\partial p$,

(8_2) $\quad d_t p = -\partial H/\partial q$,

e supponiamo H decomponibile nella somma di una hamiltoniana "principale" $H_0(t,q,p)$, in relazione alla quale possediamo un integrale completo $W_0 = W_0(t,q,P)$ della associata eq. di H.J. $H_0(t,q,\partial W_0/\partial q) + \partial W_0/\partial t = 0$ (ove le P sono le costanti da cui W_0 dipende "essenzialmente"), e di una hamiltoniana "di perturbazione" $H_1(t,q,p)$. Denotiamo con (q_0,p_0) le soluzioni delle equazioni canoniche associate alla H_0 e sotto le stesse condizioni iniziali imposte alle (q,p). Per definizione, la W_0 genera una trasformazione canonica $(t,q_0,p_0) \leftrightarrow (t,Q,P)$, per la quale le (Q,P) sono delle costanti che scriveremo (Q_0,P_0), perché la nuova hamiltoniana $K_0 =: H_0 + \partial W_0/\partial t = K_0(t,Q,P)$ (ove $Q =: \partial W_0/\partial P$) è nulla. Per la stessa trasformazione canonica generata dalla W_0: $(t,q,p) \leftrightarrow (t,Q,P)$, la nuova hamiltoniana $K = H + \partial W_0/\partial t$ è invece uguale a H_1, e le nuove variabili (Q,P) non sono costanti ma evolvono con t secondo le

(9_1) $\quad d_t Q = \partial K/\partial P = \partial H_1/\partial P$,

(9_2) $\quad d_t P = -\partial K/\partial Q = -\partial H_1/\partial Q$, nelle quali abbiamo continuato a denotare con H_1 la funzione $H_1(t,q(t,Q,P),p(t,Q,P))$ di (t,Q,P). Le (9) equivalgono a tutti gli effetti alle equazioni canoniche originali (6), e non è detto che siano più agevoli da risolvere di queste ultime, sebbene involgano la sola parte perturbativa H_1 di H. In altre parole, è equivalente risolvere le (8) sotto certe condizioni iniziali, oppure risolvere le (9) sotto le condizioni iniziali corrispondenti secondo la stessa trasformazione generata dalla W_0, per poi passare alle (q,p) mediante la trasformazione canonica inversa. Le (9) presentano tuttavia un vantaggio sulle (8) nella prospettiva di fare uso del metodo di Picard, perché per esse si possono assumere le *costanti* (Q_0,P_0) come approssimazione di ordine zero. [50] Indicando con $|_0$ la sostituzione delle (Q,P) con le (Q_0,P_0), l'approssimazione di ordine uno della soluzione si ha quindi dalle

$(9_{1,1})$ $d_tQ_1 = \partial H_1/\partial P|_0 = \partial H_1/\partial P(t,q(t,Q_0,P_0),p(t,Q_0,P_0))$,

$(9_{2,1})$ $d_tP_1 = -\partial H_1/\partial Q|_0 = -\partial H_1/\partial Q(t,q(t,Q_0,P_0),p(t,Q_0,P_0))$,

che si integrano con una quadratura sotto le condizioni iniziali $(Q_1,P_1)(t=0) = (Q_0,P_0)$. Similmente proseguendo per $n > 1$, e con analogo significato di $|_n$, abbiamo così la gerarchia di Picard di equazioni differenziali integrabili mediante una quadratura

$(9_{1,n})$ $d_tQ_{n+1} = \partial H_1/\partial P|_n$,

$(9_{2,n})$ $d_tP_{n+1} = -\,\partial H_1/\partial Q|_n$,

sotto le condizioni iniziali $(Q_n,P_n)(t=0) = (Q_0,P_0)$. È anche evidente che *se* H_1 *non dipende esplicitamente da* t, il risultato deve essere una serie di potenze di t con termine di grado zero pari a (Q_0,P_0); ad esempio, al più basso ordine significativo,

(10_1) $Q_1(t) = \partial H_1/\partial P|_0 t + Q_0$,

(10_2) $P_1(t) = -\partial H_1/\partial Q|_0 t + P_0$.

Se la successione $(Q_n(t),P_n(t))$ determinata dalle (9) converge per $n \to \infty$ ed il dato t, e se H_1 non dipende da t, il corrispondente limite $(Q(t),P(t))$ è una serie di potenze di t convergente per quel t, con termine di grado (0) pari a (Q_0,P_0).

Un esempio di applicazione del metodo di variazione delle costanti, semplice ma didatticamente efficace, è quello dell'oscillatore armonico 1-dim lungo la retta x. Supporremo unitaria la massa dell'oscillatore (e quindi k (≡ costante elastica) = ω^2 (≡ quadrato della frequenza), e per semplicità, $x(t=0) = 0$, $p(t=0) = 1$. Partendo dalla equazione di moto, sappiamo allora che $x(t) = (1/\omega)\sin(\omega t)$ e $p(t) = \cos(\omega t)$. Considereremo l'addendo cinetico $p^2/2$ della hamiltoniana $H = (p^2+\omega^2x^2)/2$ come sua parte principale H_0 e l'addendo potenziale $\omega^2x^2/2$ come sua parte di perturbazione H_1. Ciò è a prima vista un po' paradossale, ma funziona. La funzione principale di

[50] È precisamente da questa circostanza che il metodo prende il suo nome. Curiosamente, esso è spesso presentato come metodo "perturbativo", sebbene non vi si preveda alcuno sviluppo in serie di potenze di piccolo parametro.

Hamilton associata a H_0, $W_0 = W_0(t,x,P)$ ha le due determinazioni $\pm Px - P^2t/2$, ma se ne può considerare una sola senza perdita di generalità, ad esempio quella con il segno + in fronte a Px. La trasformazione canonica $(t,x,p) \leftrightarrow (t,X,P)$ è allora $P = \partial W_0/\partial x = p$, $X = \partial W_0/\partial P = x - Pt$. Le equazioni di evoluzione per queste (X,P) sono le (9), quindi nel nostro caso le

(11_1) $d_tX = \partial H_1/\partial P = \omega^2(X+Pt)t$,

(11_2) $d_tP = -\partial H_1/\partial X = -\omega^2(X+Pt)$,

per le quali è identicamente $d_tX + td_tP = 0$. Le condizioni iniziali sulle variabili (X,P) sono poi

(12) $(X(t=0),P(t=0)) = (0,1)$.

Sotto le (12), la soluzione delle (11) è

(13_1) $X(t) = (1/\omega)[\sin(\omega t) - \omega t\cos(\omega t)]$

(13_2) $P(t) = \cos(\omega t)$,

e quindi:

(14_1) $x(t) = X(t) + tP(t) = (1/\omega)\sin(\omega t)$,

(14_2) $p(t) = P(t) = \cos(\omega t)$,

come deve appunto essere. Ciò conferma l'equivalenza generale del sistema canonico per le (x,p) e di quello per le (X,P), sotto condizioni iniziali trasformate le une delle altre.

Useremo adesso l'algoritmo di Picard per (possibilmente) risolvere le (11), con le condizioni iniziali (12) come approssimazione di ordine zero, $(X_0,P_0) = (0,1)$. Poiché H_1 non contiene esplicitamente t, le risultanti (X,P)(t) devono essere serie di potenze di t aventi (0,1) come termini di grado zero. Oltre alle $(X_0,P_0) = (0,1)$, quindi $x_0(t) = X_0+P_0t = t$, $p_0(t) = 1$ (moto inerziale di momento unitario) all'ordine zero, al successivo ordine uno troviamo, eseguendo le quadrature, $X_1(t) = (1/\omega)\omega^3t^3/3$ e $P_1(t) = 1 - \omega^2t^2/2$, ossia $x_1(t) = (1/\omega)(\omega t - \omega^3t^3/3!)$ e $p_1(t) = P_1(t)$. Proseguendo in modo analogo per gli ordini successivi, otteniamo esattamente lo sviluppo in serie di potenze di t delle funzioni a 2° membro delle (13). L'analogo sviluppo in serie della funzione a 2° membro della (14_1) segue facilmente tenendo conto della identità, valida per $n \geq 1$, $1/n! - 1/[(n+1)(n-1)!] =$
$= 1/(n+1)!$.

Naturalmente il successo della procedura illustrata – nel presente caso dell'oscillatore armonico 1-dim – significa che sono soddisfatte tutte le condizioni previste dal teorema di Picard-Lindelöf. Osserviamo infine che la $W_0 = \pm Px - P^2t/2$ è realmente un integrale completo della equazione di H.J. associata alla H_0, perché risulta $\partial^2 W_0/\partial x\partial P$ ($\equiv \partial^2 W_0/\partial P\partial x$) $= \pm 1$ a seconda del segno ± in fronte a Px scelto in W_0. [51]

[51] È naturale chiedersi cosa succederebbe usando l'algoritmo di Picard *direttamente* sul sistema canonico per le (x,p) con l'hamiltoniana H, cioè sul sistema $d_tx = p$, $d_tp = -\omega^2x$, con le solite condizioni iniziali $(x,p)(t=0) = (0,1)$. Risparmiandoci di scrivere l'argomento t nei primi membri delle relazioni che seguono, si trova $x_0 = 0$; $x_1 = x_2 = t$; $x_3 =$

II.3) SULLA ANALISI FUNZIONALE-NUMERICA

Gli esempi di procedure di soluzione per approssimazioni successive che abbiamo illustrato in II.2 sono relativamente potenti e di successo, ma è difficile collocarli in un quadro organico; e del resto l'organicità resta un grosso problema in una materia così vasta e tuttora in consistente crescita quale è quella della **analisi funzionale-numerica**, o **analisi numerica astratta**. Dedicheremo quest'ultima sezione ad un esame più ravvicinato di alcuni dei suoi concetti, teoremi e tecniche generali, senza peraltro entrare nei molti specifici capitoli in cui essa viene usualmente suddivisa (applicazioni alle equazioni differenziali o integrali della fisica matematica, equazioni alle differenze finite, metodo degli elementi finiti, metodi spettrali, teoria degli operatori monotoni, teoria delle perturbazioni, ecc.).

A nostro avviso, per il momento nemmeno i più ambiziosi tentativi didattici riescono a dare un'idea ragionevolmente unitaria della teoria analitica dell'approssimazione. Coniugata con la fisica teorica, essa ha prodotto quell'ampio campo di "attività di calcolo" che va oggi sotto il nome di **fisica computazionale**, e il cui ruolo centrale nelle scienze esatte e nella ingegneria contemporanee non ha bisogno di essere sottolineato. La fusione di concetti e metodi propri dell'analisi funzionale con quelli della analisi numerica classica risale all'inizio dello scorso secolo, ma un efficace interfacciamento tra teoria ed applicazioni concrete si è reso possibile soltanto a partire dagli anni '50 circa, con l'avvento dei calcolatori elettronici programmabili. Da allora, la scienza computazionale è stata sfidata a rimodellarsi da una parte sulla nuova disciplina matematica – appunto, l'analisi funzionale-numerica – e dall'altra sulla disponibilità di sempre più potenti strumenti di calcolo automatico.

Si può discutere sui modi e sulla misura in cui questa sfida è stata realmente raccolta. Il ruolo fondamentale del calcolatore a supporto dell'analisi (e talvolta addirittura dell'euristica) dei modelli teorici, tipicamente equazionali, della scienza contemporanea è fuori questione; ma si deve tener presente che nella maggior parte dei casi, e specialmente in rapporto ai modelli più complessi, gli algoritmi di calcolo usati nella pratica quotidiana *non sono fondati*. Vale a dire, di norma mancano teoremi che attestino l'esistenza e/o l'unicità delle soluzioni ricercate; non è provato che le successioni di soluzioni approssimate da essi prodotte convergano in qualche modo alle effettive soluzioni (o semplicemente convergano a qualcosa) quando i parametri di accuratezza tendono ai loro limiti naturali; non sono disponibili stime a priori degli errori relativi o della velocità della

---

$= x_4 = (1/\omega)(\omega t - \omega^3 t^3/3!)$; …; $p_0 = p_1 = 1$; $p_2 = p_3 = 1 - \omega^2 t^2/2!$; $p_4 = p_5 = 1 - \omega^2 t^2/2! + \omega^4 t^4/4!$; … Le serie sono pertanto quelle corrette, ma la loro costruzione procede "a velocità dimezzata". Dopotutto l'algoritmo di Picard si fa dunque onore, in questo caso, anche ignorando il sistema canonico (9), ovvero la trasformazione canonica generata da W_0, integrale completo dell'equazione di H.J. associata alla parte principale di H.

possibile convergenza, o dimostrazioni della stabilità delle procedure, … e così via. D'altra parte è cosa ovvia che il calcolo su elaboratore, per quanto grandi possano essere i suoi benefici pratici, ha ben poco da offrire ad una vera e propria "analisi" del modello fenomenologico considerato se non è supportato da algoritmi fondati: perché tutto si riduce in tal caso ad una mera "sperimentazione mediante calcolatore", essenzialmente pilotata dall'intuizione e dalla esperienza pratica.

La distinzione tra calcolo mediante algoritmi *fondati* e calcolo mediante algoritmi *non fondati* suggerisce poi una analoga distinzione, nei fatti largamente acquisita, tra una fisica computazionale "razionale" e una fisica computazionale "sperimentale" (per così dirle); quella razionale (sorella povera dell'altra) momento di genuina analisi del modello, e quella sperimentale per lo più orientata alle applicazioni più richieste e immediatamente remunerative. [52] Nell'opinione di alcuni matematici applicati contemporanei, è addirittura possibile che nel giro di qualche decina di anni il baricentro dell'attività matematica come la si è fin qui concepita si sposterà del tutto sui problemi legati al progresso degli algoritmi, alla loro ottimizzazione in termini di rapporto costi/benefici, alla ricerca operativa, all'analisi combinatoria su grandi insiemi, al controllo ottimale, alla statistica, .. e via dicendo. Questo fenomeno, in cui spesso la forza bruta spiegabilmente prevale sull'alternativa di esercitare la propria destrezza nel modo tradizionale, è infatti pienamente in atto; ed il calcolatore, che ne è stato l'impulso scatenante, gli continua a fornire in sovrabbondanza l'energia necessaria. [53] Si conclude così che anche quel personaggio che abbiamo più sopra introdotto come "risolvitore di equazioni", si scinde ormai in due figure professionali *ben distinte*: quella dello specialista di analisi numerica astratta, e quella dello sperimentatore mediante computer, tra loro da sempre in precario e insufficiente contatto cooperativo.

La vastità della materia obbliga a limitarci ad una piccola scelta di problemi fondamentali e molto generali. Torniamo alla nozione di equazione (generalmente non lineare), nella forma

(1) $A(x) = 0,$

ove A: $D(A)$ ($\subset X$) $\to Y$, $D(A)$ è al solito il dominio di definizione di A, X e Y sono B-spazi (il secondo con zero $O_Y \equiv 0$) e $x \in D(A)$ è l'incognita. Se in particolare $Y \equiv X$, $A(x)$ può essere sempre scritta come $B(x) - x$ ponendo $B(x) =: A(x) + Id_X x$; la (1) diventa allora

[52] È qui opportuno ricordare come primo e storico esempio di fisica computazionale *sperimentale* la pionieristica indagine realizzata da E. Fermi, J. Pasta e S. Ulam ("Studies in Nonlinear Problems", Los Alamos Report LA 1940, 1955) sul comportamento di una catena di oscillatori debolmente accoppiati in modo anarmonico: modello noto da allora come "FPU model", e pietra miliare nell'approccio numerico ai problemi della dinamica non lineare di sistemi hamiltoniani a molti gradi di libertà (vedi ad es. in "E. Fermi Collected Papers", ed. E. Segré, Un. of Chicago Pr. (1965), vol 2, p. 978). I risultati furono «francamente sorprendenti» (Ulam, in "Adventures of a Mathematician", 1983), nel senso che l'attesa equipartizione dell'energia dei modi del sistema non appariva. Soltanto molto più tardi si chiarì che questa assenza era dovuta alla scelta di valori troppo bassi dell'energia, situati nel cosiddetto "slow mixing regime".

[53] Non ci sembra molto sensato deplorare più di un tanto la situazione, che è nella forza dei fatti. Si legga ad esempio il saggio dal tono un po' apocalittico di C. Truesdell: "Il calcolatore: rovina della scienza e minaccia per il genere umano" in "La nuova ragione. Scienza e cultura nella società contemporanea", Bologna, Scientia - Il Mulino, (1980).

(2) $B(x) = x$,

che è evidentemente una equazione di punto fisso (vedi II.2) con l'incognita x in D(B) (≡ dominio di B) ≡ D(A). È certamente motivo di sorpresa, per il non-specialista, constatare quanta parte della moderna analisi funzionale verta su problemi di punto fisso. [54]

Riferendoci alla (2), sia $B: Q \subset X \to X$ (≡ B-spazio) tale che esista una costante $q \in (0,1)$ per cui $\|B(x')-B(x'')\|_X \le q\|x' - x''\|_X$ per ogni coppia x', x'' in Q. In questo caso B applica Q in Q e si dice una **contrazione su** Q **con rapporto** (di contrazione) q. Abbiamo allora il fondamentale **teorema di S. Banach** (1921) sulla (2):

T_1). «Se $Q \subset X$ è chiuso, e B è una contrazione su Q con rapporto q, l'equazione (2) ha una e una sola soluzione $\underline{x} \in Q$ che è il limite per $n \to \infty$ della successione $\{x_n\}_{n=0,1\,\ldots}$ generata dalla formula ricorsiva $x_{n+1} = B(x_n)$, ogni elemento della quale è in Q se lo è x_0. La stima a priori dell'errore (alla n-ma iterazione) è data da

(3) $\|x_n - \underline{x}\|_X \le \|B(x_0) - x_0\|_X\, q^n/(1-q)$».

Si apprezzi la potenza di (T_1), che sotto le ipotesi menzionate fornisce simultaneamente una prova di esistenza/unicità della soluzione, una procedura effettiva per la sua costruzione e una stima a priori dell'errore. Di (T_1) esistono parecchie interessanti applicazioni e generalizzazioni. Ad esempio la teoria di Picard per la soluzione della equazione differenziale (II.2, 6,6′), ormai interpretata in senso astratto con x(t) in un B-spazio e con t-derivata nel senso di Fréchet, è riconducibile (si può dimostrare) al teorema di Banach.

Passando alla più generale equazione (1), molto importante è il **teorema di Kantarovich**, (o **di Newton-Kantarovich**, 1948 [55]), che generalizza in direzione astratta il ben noto "metodo delle tangenti" proposto da Newton circa due secoli e mezzo prima:

T_2). «Ipotesi: (a): $A: D(A)\ (\subset X) \to Y$ è Fréchet-derivabile nella palla (aperta) di D(A) di centro x_0 e raggio r, $\mathcal{B}_r(x_0)$, e la sua F-derivata A' è ivi Lipschitz-continua con costante (di Lipschitz) $c > 0$; (b): in $\mathcal{B}_r(x_0)$, A' è continuamente (uniformemente) invertibile e l'inversa $(A')^{-1}$ ha norma [56] non maggiore di $m > 0$; (c): esiste un $\eta \ge \|A(x_0)\|_Y$ tale che $q =: \eta m^2 c/2 < 1$ e $r' =: \eta m \sum_{h=0}^{\infty} q^{\alpha(h)} < r$, ove $\alpha(h) =: 2^h - 1$; Tesi: (d): l'equazione (1) ha una soluzione $\underline{x} \in [\mathcal{B}_{r'}(x_0)]$ (dove [ ] denota la chiusura, qui in X) che è il limite per $n \to \infty$ della sequenza $\{x_n\}_{n=0,1\,\ldots}$ (anch'essa in $[\mathcal{B}_{r'}(x_0)]$), generata dalla formula ricorsiva $x_{n+1} = x_n - (A'(x_n))^{-1}A(x_n)$; (e): la stima a priori dell'errore è data da

---

[54] Ad esempio il trattato di analisi funzionale di E. Zeidler (1985) dedica all'argomento il primo dei suoi cinque volumi, di circa 900 pagine.

[55] Vedi in particolare L.V. Kantarovich, G.P. Akilov, "Functional Analysis in Normed Spaces" (engl. transl.) Pergamon Pr. 1964, p. 658.

[56] Si ricorda che la norma ($\|\cdot\|$) di un operatore lineare è il sup della norma dei suoi valori sulla sfera unitaria (supposto esistere finito).

(4) $\quad ||x_n - \underline{x}||_X \le \eta m q^{\alpha(n)}/(1-q^{\beta(n)})$,

ove $\beta(n) =: 2^n$.»

Sotto le ipotesi (a÷c), anche il teorema di Kantarovich offre dunque una prova costruttiva dell'esistenza di una soluzione, una procedura effettiva per il suo calcolo, ed una stima a priori dell'errore. Il problema dell'unicità non è invece considerato nella versione qui riportata (ne esistono altre).

In generale, l'analisi funzionale fornisce sia (i) teoremi di esistenza *costruttivi*, come (T_1) e (T_2), sia (ii) teoremi di esistenza *non costruttivi*, sia infine (iii) metodi di soluzione *costruttivi* eventualmente supportati da un'ipotesi o da un asserto di esistenza. In certi casi, unendo le risorse dei risultati delle classi (ii, iii), nonché la necessaria dose di elaborazione matematica, si sono ottenuti algoritmi *fondati* per la soluzione dell'equazione di partenza, per lo più di tipo (1).

Cominceremo col presentare alcuni esempi della classe (ii) di teoremi, che hanno tutti comune natura topologica. Un prototipo di immediata evidenza intuitiva, risalente a B. Bolzano (1817) ne è il seguente:

T_3). «Se f: $[0,1] \to \mathbf{R}$ è continua, e $f(0)f(1) < 0$, allora l'equazione $f(x) = 0$ ha (almeno) una soluzione in $(0,1)$.»

La dimostrazione di (T_3) riposa sul fatto, a sua volta provato in qualunque trattato di analisi istituzionale, che una successione di reali non decrescente [non crescente] e limitata verso l'alto [verso il basso] ha un limite finito.

Molto meno intuitivo è il **teorema di L. Brouwer** (1912), che recita:

T_4). «Una trasformazione continua di una palla chiusa di $\mathbf{R}^{n\ge1}$ in se stessa ha ivi un punto fisso». (Qui e nel seguito "un punto fisso" sta ovviamente per "almeno un punto fisso"). In un linguaggio più suggestivo, (T_4) afferma che se un continuo materiale (ad esempio un liquido) che riempie una palla viene mescolato arbitrariamente sotto il vincolo di subire uno spostamento *continuo*, alla fine c'è una sua particella che *non si è spostata*. Di (T_4) esistono versioni più generali ed astratte, ad esempio:

T_5). «sia Ω è un sottoinsieme chiuso e convesso di un B-spazio n-dimensionale, e $A : \Omega \to \Omega$ una applicazione continua; allora A ha un punto fisso in Ω.»

Nella nostra opinione, nessuna delle dimostrazioni correnti del teorema di Brouwer è abbastanza semplice per essere accennata qui.

Proseguiamo con il **teorema di J. Schauder** (1930):

T_6). «sia Ω un sottoinsieme limitato, chiuso e convesso di un B-spazio e $A : \Omega \to \Omega$ un'applicazione compatta. [57] Allora A ha un punto fisso in Ω ».
La dimostrazione standard fa uso del teorema di Brouwer.

Tra le conseguenze più note del teorema di Schauder ricordiamo ancora:
(T_7). «sia $[\mathcal{B}_r(O_X)]$ la palla chiusa di centro O_X e raggio r del B-spazio X, e sia A: $[\mathcal{B}_r(O_X)] \to X$ compatta. "Se $A(x) \neq \lambda x$ per ogni $\lambda > 1$ e per ogni $x \in X$ di norma r", A ha un punto fisso in $[\mathcal{B}_r(O_X)]$. La stessa conclusione vale se l'ipotesi riportata tra " " è sostituita dalla "se $\|A(x)\|_X \leq r$ per ogni $x \in X$ di norma r, $\|x\|_X = r$".

Concludiamo questa lista di teoremi di punto fisso con il **teorema di J. Leray e Schauder** (1934), che recita:
T_8). «sia f: $X \to X$ (X essendo un B-spazio) compatta, ed esista un $C > 0$ tale che, per ogni $x \in X$ soddisfacente alla $tf(x) = x$ con t in $[0,1]$, risulti $\|x\|_X \leq C$. Allora f ha un punto fisso in X».

Venendo ai metodi costruttivi di soluzione di una equazione astratta (generalmente tra B-spazi), essi hanno più o meno la seguente comune natura: precisato il tipo di equazione che si intende risolvere per successive approssimazioni, si comincia con l'"approssimarla" mediante una successione di convenienti equazioni facilmente risolvibili. L'obiettivo consiste allora nel provare che la successione delle soluzioni approssimate così ottenute converge, sotto le ipotesi del caso e nel senso opportuno, alla soluzione della equazione di partenza. Quando è possibile, si fornisce una stima a priori dell'errore e del tasso di convergenza. Spesso la procedura dà al contempo un teorema di esistenza, ed è allora completamente autonoma. Esiste una impressionante costellazione di variazioni su questo tema, spesso collegate "geneticamente" tra loro. Un naturale intento didattico, che ovviamente non può essere perseguito qui, dovrebbe essere quello di presentare tali procedure "per filiere" di problemi analoghi e di generalità crescente (o decrescente, a seconda dei gusti), evidenziando i collegamenti tra caso e caso; ma è raro trovare ispirazioni del genere nella pur vasta letteratura specializzata. Ci limiteremo dunque a qualche esempio in tal senso, beninteso non proponendoci nulla più che di offrire al lettore il "sapore" della complessa materia.

Il caso più semplice è probabilmente quello del cosiddetto **metodo dei momenti**. Riferendoci ad una equazione del tipo (1), la soluzione approssimata di ordine (n) viene espressa come combinazione lineare di n elementi di una data base $\{\phi_i\}$ di X (B-spazio includente D(A)) secondo la $x_n = \sum_1^n c_i\phi_i$. Se $\{\psi_j\}$ è una simile base del duale Y* di Y (B-spazio includente R(A)), si ottiene così

(5_n) $(A(\sum_1^n c_i\phi_i), \psi_j) = 0$,

[57] Ricordiamo che una funzione f: $X \to Y$, ove X e Y sono B-spazi, si dice compatta (in X) se è ivi continua e la

per j = 1, .., n, dove con (f,g) si è denotata la forma bilineare uguale al valore del secondo argomento g (un funzionale lineare limitato sul B-spazio Y, elemento del suo duale Y*) nel primo argomento f (un elemento del B-spazio Y). Le (5_n) sono n equazioni (generalmente non lineari) negli n coefficienti $c_{i=1, .., n}$. x_n è determinato se esiste una soluzione unica delle (5n); ma fino a quando la situazione rimane così generale, non è detto che la successione $\{x_n\}$ converga in X per n $\to \infty$, e tanto meno che converga ad una soluzione della (1).

Supporremo allora che X e Y siano spazi di Hilbert (H-spazi), e che la (1) sia del tipo

(6) $Lx - y = 0$

(denotiamo al solito con 0 lo zero di Y), ove L: D(L) $\subset$ X $\to$ Y è una applicazione *lineare* (non necessariamente limitata) e y è dato in Y. Diamo per noto che, essendo Y un H-spazio, Y e Y* possono essere identificati, sussistendo tra essi una isometria lineare-duale che lascia invariate le norme ($\|\ \|_{Y*} = \|\ \|_Y$), e secondo la quale la precedente forma bilineare si identifica con il prodotto interno standard in Y. In questo caso la base $\{\psi_j\}$ è in Y e le (5_n) diventano

(5_nbis) $\sum_1^n (L\phi_i,\psi_j)c_i = (y,\psi_j)$

per j = 1, .., n, e dove (,) è il prodotto interno in Y. Le (5_nbis) sono un sistema di n equazioni lineari negli altrettanti coefficienti c_i, la cui risolvibilità unica è espressa dalla usuale condizione $\det((L\phi_i,\psi_j)) \neq 0$.

Siano poi $\{X_n\}$ e $\{Y_n\}$ successioni di sottospazi di X e rispettivamente di Y, la prima tale che $X_n \subset D(A)$ $\forall n$, e sia $\{P_n\}$ una successione di proiezioni che applicano Y su Y_n, quindi $P_n^2 = P_n$ su Y, e $P_nY = Y_n$. Si consideri la successione di equazioni lineari approssimanti

(7_n) $P_n(Lx_n-y) = 0$,

con $x_n \in X_n$. Se le P_n sono ortoproiezioni, e si identificano X_n e Y_n con le varietà lineari abbracciate da $\{\phi_i\}_1^n$ e rispettivamente da $\{\psi_j\}_1^n$, si verifica facilmente che la (5_nbis) e la (7_n) si equivalgono. Tuttavia anche in assenza di questa equivalenza, e ferme restando le $X_n \subset D(A)$, $P_n^2 = P_n$, $P_nY = Y_n$, la (7_n) conserva il significato di equazione approssimante la (6). Il metodo di soluzione della (6) facente capo alla (7_n) si dice **metodo delle proiezioni** (per equazioni lineari). La sua giustificazione è contenuta nel seguente teorema di convergenza.

T_9) Ipotesi: (a): il dominio D(L) e il range R(L) sono densi in X e rispettivamente in Y; (b): L: D(L) $\to$ R(L) è biiettiva; (c): X_n e Y_n hanno dimensione finita ed uguale (tipicamente la loro dimensione è n); (d): le proiezioni P_n sono uniformemente limitate, cioè esiste una costante C > 0 per cui $\|P_n\| \le C$ $\forall n$; (e): $\forall y \in Y$, la distanza $d(y,LX_n)$ di y da LX_n ($\equiv \inf_{z\in LXn}\|z-y\|_Y$) tende a zero per n $\to \infty$; (f): avendo posto $\tau_n =: \inf_{z\in LXn,\ \|z\|=1}\|P_nz\|_Y$, risulta $\underline{\lim}_{n\to\infty}\tau_n > 0$. Tesi: (g): (almeno) da

---

chiusura di f(σ), inclusa in Y, è compatta per ogni sottoinsieme limitato σ di X.

un dato n in su, per ogni $y \in Y$ la (7_n) ha un'unica soluzione x_n, e $||Lx_n-y||_Y \to 0$ per $n \to \infty$; (h): il tasso di convergenza di Lx_n a y è stimato dalla $||Lx_n-y||_Y \leq (1+C/\tau_n)d(y,LX_n)$. [58] #

Sotto le ipotesi (a÷f), il metodo delle proiezioni per equazioni lineari fornisce dunque anche una prova di esistenza (per qualunque y). Il suo limite principale sta nell'ipotesi (d) (che è spesso difficile soddisfare), e naturalmente nella linearità della equazione di partenza (6).

Come ultimo esempio illustreremo ora uno dei cosiddetti "schemi di approssimazione" per la soluzione di una equazione generalmente non lineare del tipo (1). La nostra scelta è caduta su di esso sia per la sua generalità e potenza, sia perché il teorema sul quale si fonda si può dimostrare con mezzi abbastanza elementari. L'idea di base è ancora quella di "approssimare" la (1) con una successione di equazioni (generalmente non lineari) del tipo

(8_n) $A_n(x_n) = O_{Yn}$,

n = 1, 2 .., ove A_n: $D(A_n)$ ($\subset X_n$) $\to Y_n$, X_n e Y_n essendo B-spazi, il secondo con zero O_{Yn}, sotto appropriate condizioni, in modo che la successione delle soluzioni delle (8_n) "approssimi", nel senso che preciseremo, una soluzione della (1).

Ci serviranno parecchie definizioni, che comunque hanno valore anche di per sé. Sia K un B-spazio, e $\{K_n \supset K\}_{n=1, \dots}$ una successione di B-spazi includenti K. Sia poi $\{Z_n\}_{n=1, \dots} \equiv \{Z_n\}$ (qui e nel seguito in questi casi aboliremo per brevità il pedice $_{n=1, \dots}$) una successione di B-spazi e $\{W_n\}$ una successione di *suriezioni lineari continue* [59] W_n: $K_n \to_{sur} Z_n$, quindi $Z_n = W_nK_n$. La $\{W_n\}$ come descritta si dice uno **schema di restrizione su** K (infatti le W_n non sono in generale iniettive, e quindi "restringono" i B-spazi K_n in corrispondenti Z_n, di cardinalità non maggiore). La disponibilità di $\{W_n\}$ permette di definire un particolare tipo di convergenza di una successione $\{z_n \in Z_n\}$ ad un elemento $z \in K$. Precisamente, diremo che $\{z_n\}$ W-**converge a** z (o che z è un W-**limite di** $\{z_n\}$) se

(9) $||z_n-W_nz||_{Zn} \to 0$

per $n \to \infty$. [60] Non è detto che $\{z_n\}$ abbia sempre un W-limite in K; o che se ne ha, ne abbia uno solo. Se ogni $z \in K$ è un W-limite di qualche $\{z_n\}$, diremo che la successione dei B-spazi $\{Z_n\}$ W-**converge a** K. D'altra parte $\{z_n\}$ *ha al più* un W-limite in K sse, $\forall z \in K$,

(10) $(\lim_{n\to\infty}||W_nz||_{Zn} = 0) \Rightarrow (z = 0)$. [61]

---

[58] Per la dimostrazione, si veda ad es. in M.A. Krasnoselskii et al., loc. cit. nella Bibliografia Generale.

[59] Ricordiamo che la continuità di una applicazione lineare è equivalente alla sua limitatezza. È ovvio che il trasformato di un B-spazio sotto una tale applicazione è ancora un B-spazio.

[60] La W-convergenza di $\{z_n\}$ a z include come caso particolare l'usuale convergenza in norma in K: basta porre $K_n = K$ e $W_n = Id_K$, quindi $Z_n = K$, perché la (6) diventi $||z_n-z||_K \to 0$.

[61] 1): La (10) è *sufficiente* all'unicità: supposto infatti che $||z_n-W_nz||_{Zn} \to 0$ valga per $z = z'$ e per $z = z''$, per la la disuguaglianza triangolare si ha $||W_n(z'-z'')||_{Zn} \to 0$, e quindi $z' = z''$ in forza della (10). 2): La (10) è conseguenza

La (10) si dice **condizione di** W**-non-degenerazione delle norme** $\|\ \ \|_{Z_n}$.

Siano adesso $\{T_n\}$ e $\{V_n\}$ schemi di restrizione su G e rispettivamente su H, secondo le

(11_n) T_n: $G_n \to_{sur} X_n$, $X_n = T_nG_n$, $G \subset G_n$, e

(12_n) V_n: $H_n \to_{sur} Y_n$, $Y_n = V_nH_n$, $H \subset H_n$,

ove G_n, X_n, G, H_n, Y_n e H sono tutti B-spazi. Tornando alla equazione (generalmente non lineare) (1), ove A: D(A) ($\subset$X) $\to$ Y, con $O_Y \in$ R(A) $\equiv$ range di A (in caso contrario la (1) non avrebbe soluzioni), supporremo che $\forall$n, $G_n \subset X$ e $H_n \subset Y$, e inoltre che D(A) $\subset$ G e R(A) $\subset$ H. Sia poi $\{A_n\}$ una successione di applicazioni (generalmente non lineari) del tipo

(13_n) $A_n : D(A_n) \subset X_n \to Y_n$.

Una successione di applicazioni $\{A_n\}$ come la (13_n), ogni elemento della quale è definito in un dominio incluso in un B-spazio e prende i suoi valori in un B-spazio, si dice **stabile** se esiste una funzione $\omega = \omega(t\in[0,\infty])$ con valori in $[0,\infty]$ continua e strettamente crescente tra 0 (ove $\omega(0) = 0$) e ∞ (ove $\omega(\infty) = \infty$) e tale che, $\forall(x_n', x_n'') \in D(A_n)$ e per ogni n (o almeno da un certo n in su), risulti

(14_n) $\omega(\|x_n'-x_n''\|_{X_n}) \le \|A_n(x_n')-A_n(x_n'')\|_{Y_n}$.

Supponiamo ora che per la successione $\{A_n\}$ sussista il vincolo:

(15_n) $T_nD(A) \subset D(A_n)$

$\forall n \ge 1$. Allora $\forall x \in D(A)$ sono definiti T_nx ($\in X_n$) e $A_n(T_nx)$ ($\in Y_n$), e inoltre A(x) ($\in$R(A) $\subset$ H) e $V_nA(x)$ ($\in Y_n$), nonché la differenza $A_n(T_nx) - V_nA(x)$ ($\in Y_n$). Si supponga a questo punto che, per un dato $x \in D(A)$, sia

(16) $\|A_n(T_nx)-V_nA(x)\| \to 0$

per $n \to \infty$. Le (15_n) e la (16) si dicono **condizioni di** (T,V)**-approssimazione** (la seconda **in** x) **per la successione** $\{A_n\}$; se esse sono soddisfatte, la successione $\{A_n\}$ si dice uno **schema di** (T,V)-**approssimazione di** A **in** x. Se $\{A_n\}$ è un tale schema in una soluzione della (1), una soluzione della (8_n) (che per definizione $\in D(A_n) \subset X_n$) si dice una **soluzione approssimata di ordine** (n) della (1) stessa (rispetto allo schema $\{A_n\}$).

La (16) asserisce che la successione $A_n(T_nx)$ V-converge a $A(x) \in R(A)$ ($\subset$ H). Essa può rinforzarsi come segue. Si supponga che esistano una costante $k > 0$ e una funzione $\alpha(x) > 0$ tali che, per il dato $x \in D(A)$, e per ogni n (o almeno da un certo n in su)

(17) $\|A_n(T_nx) - V_nA(x)\|_{Y_n} \le \alpha(x)n^{-k}$,

cioè che il 1° membro della (14) sia $O(n^{-k})$ per $n \to \infty$; allora lo schema di (T,V)-approssimazione $\{A_n\}$ (di A in x) si dice brevemente **di ordine** k. È chiaro che la (17) implica la (16).

---

necessaria dell'unicità: infatti la successione $\{W_nz\}$ W-converge a z, e inoltre W-converge anche a 0 in forza

Se poi $\underline{x} \in D(A)$ è soluzione della (1), la (16) e la (17) diventano, per $x = \underline{x}$,

(16bis) $\quad ||A_n(T_n\underline{x})||_{Y_n} \to 0$

per $n \to \infty$, e rispettivamente

($17bis_n$) $\quad ||A_n(T_n\underline{x})||_{Y_n} \le \alpha(x)n^{-k}$

da un certo valore di n in su.

Possiamo finalmente enunciare il teorema in questione.

T_{10}). «Per la data (1), dati schemi di restrizione $\{T_n\}$ (su $G \subset X$) e $\{V_n\}$ (su $H \subset Y$) e dato schema di (T,V)-approssimazione $\{A_n\}$ descritti come sopra, abbiamo: Ipotesi: (a): le norme $||\ \ ||_{X_n}$ sono (T)-non-degeneri; (b): l'eq. (1) ha (almeno) una soluzione esatta e una soluzione approssimata di ogni ordine; (c): valgono le condizioni di approssimazione (15_n) e (16bis), quest'ultima in ogni soluzione della (1); (d) la successione $\{A_n\}$ è stabile. Tesi: (e): La soluzione $\underline{x}$ della (1) è unica, e la soluzione $\underline{x}_n$ della (8_n) è unica almeno da un certo valore di n in su; (f) la successione $\{\underline{x}_n\}$ T-converge a $\underline{x}$; (g) se $\{A_n\}$ è $O(n^{-k})$ in $\underline{x}$ (soluzione della (1)), e $\omega(t) \sim \gamma t^{\rho}$ (con γ, $\rho > 0$) per $t \to 0+$, allora $||\underline{x}_n - T_n\underline{x}||_{X_n} = O(n^{-k/\rho})$, cioè $\underline{x}_n$ T-converge a $\underline{x}$ come $n^{-k/\rho}$ converge a zero.»

Dim. Tesi (e): siano x', x'' due soluzioni della (1) (ipotesi (b)). Poiché la successione $\{A_n\}$ è stabile (ipotesi (d)), introducendo la funzione inversa ω^{-1} della ω, al pari di quest'ultima non decrescente e continua tra gli stessi estremi, abbiamo $||T_n(x'-x'')||_{X_n} \le \omega^{-1}(||A_n(T_nx')-A_n(T_nx'')||_{Y_n}) \le \omega^{-1}(||A_n(T_nx')||_{Y_n}) + \omega^{-1}(||A_n(T_nx'')||_{Y_n})$ (disuguaglianza triangolare). L'ultimo membro $\to 0$ per $n \to \infty$ in forza della (16) (ipotesi (c)), quindi $||T_n(x'-x'')||_{X_n} \to 0$, e in conclusione $x' = x''$ in forza dell'ipotesi (a). Siano poi x_n' e x''_n due soluzioni della (8_n) (ipotesi (b)); tornando a fruire della stabilità della $\{A_n\}$, $\omega(||x_n'-x_n''||_{X_n} \le 0$, e quindi $x_n' = x_n''$. Tesi (f): ancora partendo dalla stabilità di $\{A_n\}$ e denotando con $\underline{x}$ e con $\underline{x}_n$ la soluzione della (1) e rispettivamente della (8_n), abbiamo $||x_n-T_n\underline{x}||_{X_n} \le \omega^{-1}(||A_n(x_n)-A_n(T_n\underline{x})||_{Y_n}) = \omega^{-1}(||A_n(T_n\underline{x})||_{Y_n})$. Qui l'argomento di ω^{-1} tende a 0 in forza della (16bis) (ipotesi (c)) e per la continuità di ω^{-1} in 0, $||x_n-T_n\underline{x}||_{X_n} \to 0$, qed. La tesi (g) è un corollario di questi risultati, e la sua verifica è lasciata al lettore.#

In effetti la tesi (f) afferma soltanto che $\{\underline{x}_n\}$ T-*converge a* $\underline{x}$, e tale T-convergenza *non* implica in generale la convergenza standard. Una condizione sufficiente sotto la quale le due convergenze ad un generico x si equivalgono è quella cosiddetta **di regolarità di** T **in** x: avendo presupposto $G_n = X$ e X_n ($= T_nX$) $\subset X$ per ogni n, si assume cioè che $||T_nx-x||_X \to 0$ per $n \to \infty$. La dimostrazione è immediata, e si basa sull'uso della disuguaglianza triangolare. [62] Benché siano suggestive in tal senso, le ipotesi di (T_{10}) non sono sufficienti a passare dalla T-convergenza alla

dell'antecedente della (10). Ma allora $z = 0$ per l'ipotesi di unicità del W-limite, e quindi vale la (10).

convergenza (di $\{\underline{x}_n\}$ a $\underline{x}$); tuttavia sotto convenienti restrizioni addizionali (come quella di regolarità nella soluzione $\underline{x}$, o in generale in D(A)) si possono fondare su (T_{10}) metodi costruttivi di soluzione della (1). Alcuni **schemi alle differenze** per la soluzione di equazioni differenziali ordinarie o alle derivate parziali rientrano appunto, sotto le convenienti restrizioni, nello schema di approssimazione descritto. È il caso del notissimo schema alle differenze di Eulero (anche nella sua versione cosiddetta "accelerata") per la soluzione della generica equazione differenziale del 1° ordine normale, generalmente non lineare,

(18) $dx/dt + f(t,x(t)) = 0$,

con $0 \le t \le T$ e f continua, sotto la condizione iniziale

(18′) $x(0) = 0$.

Lo schema di Eulero, ricordiamo, è dato da $(x_k-x_{k-1})/\tau + f(t_{k-1}, x_{k-1}) = 0$, ove $k = 1,2, .. n$, $\tau = T/n$, $x_0 = 0$, $t_0 = 0$, $t_{1\le k\le n} = k\tau$. Nello schema accelerato si ha invece $(x_k-x_{k-1})/\tau + [f(t_{k-1}, x_{k-1}) + f(t_k,x_k)]/2 = 0$. Lo schema accelerato è evidentemente "implicito", cioè non immediatamente risolvibile rispetto a x_k come il precedente. Nello schema di Eulero, X è l'insieme dei grafi funzionali $x = \langle t,x(t)\rangle$ su $[0,T]$, $G_n = X\ \forall n \ge 1$, e X_n è l'insieme dei grafi funzionali $x_{(n)} = \langle t_k,x_k\rangle_{k=1}^{k=n}$ sulla maglia $\Delta_n =: \{t_0=0,t_{1\le k\le n-1}=\tau k,t_n=T\}$, quindi $X_n \subset X$, sempre $\forall n \ge 1$. Lo schema di restrizione $T_n : X \to_{sur} X_n = T_nX$ è definito da $T_nx = \langle t_k,x(t_k)\rangle_{k=1}^{k=n} \in X_n$, $\forall x \in X$. È immediato vedere che questa scelta garantisce la condizione di regolarità in x usando la norma $\|\cdot\| = \max_{[0,T]}|\cdot|$. Quanto allo schema di approssimazione, nello schema di Eulero semplice si prende $[(x_k-x_{k-1})/\tau + f(t_{k-1}, x_{k-1})]_{k=1}^{k=n}$ (colonna a n termini) per $A_n(x_{(n)})$, e dunque

(19_n) $A_n(T_nx) = [(x(t_k)-x(t_{k-1}))/\tau + f(t_{k-1}, x(t_{k-1})]_{k=1}^{k=n} = [(x(t_k)-x(t_{k-1}))/\tau - dx/dt(t_{k-1})]_{k=1}^{k=n}$;

la condizione di approssimazione è così $\max_{1\le k\le n} |(x(t_k)-x(t_{k-1}))/\tau - dx/dt(t_{k-1})| \to 0$ per $n \to \infty$. Si può anche dimostrare che la condizione di stabilità è soddisfatta, e si conclude che lo schema è convergente. Una semplice modifica permette di trattare in modo analogo la stessa equazione con la condizione iniziale $x(0) = a \in \mathbf{R}$.

Abbiamo con ciò concluso la nostra rassegna esemplificativa minimale di metodi computazionali basati su algoritmi fondati per la soluzione approssimata di equazioni in B-spazi. Ci aspettiamo che essi siano apparsi alquanto "formali", o addirittura esotici, agli occhi del non specialista; e almeno in parte, ciò è certamente dovuto all'averne proposto versioni molto generali. Ma di fatto, quei metodi di calcolo fondati (che in genere hanno almeno mezzo secolo di vita), e i molti altri con i quali avremmo potuto allargare la nostra rassegna, *sono tutti di assai concreta sostanza*. Più legittima sarebbe pertanto la perplessità di chi si chiedesse quante di tali procedure

[62] Da una parte si ha infatti $\|x_n-x\| \le \|x_n-T_nx\| + \|T_nx-x\|$; e dall'altra, $\|x_n-T_nx\| \le \|x_n-x\| + \|x-T_nx\|$. (Tutte le norme sono su X).

siano state o siano correntemente impiegate nella pratica calcolatoria quotidiana, a partire dall'avvento del computer. La risposta è già stata ampiamente suggerita. Sorretto dallo spettacolare aumento della potenza dell'hardware, e spesso giustificato dalla obiettiva complessità dei problemi, l'approccio dominante, nell'analisi computazionale dei modelli fenomenologici, rimane marcatamente empirico e pragmatico: seppure il problema è preso in considerazione, nella grande maggioranza dei casi la possibilità di lavorare con algoritmi fondati è vista cioè come una specie di necessità o garanzia di principio, alla quale si può ben rinunciare in favore di preoccupazioni più pressanti e di attività più fruttuose. Che questo sia dovuto a difficoltà obiettive, alla sottostima dell'importanza di quella garanzia, alla reale ridondanza delle dimostrazioni, a confidenza nelle risorse dell'intuizione (o della buona fortuna), a pigrizia, a pura ignoranza, o alle varie combinazioni di questi ed altri effetti, dipende dalle circostanze specifiche.

Una situazione particolarmente critica è quella che si produce quando il modello equazionale in oggetto è malsano (vedi I.1), cioè quando l'equazione che esso genera *non ha soluzioni*, almeno dove le si ricerca. La formulazione del modello equazionale dovrà essere allora "indebolita minimalmente". L'esperienza ci insegna che questa operazione, se possibile, implica una forte dose di creatività matematica, e può aprire prospettive di notevole portata. La moderna teoria delle "soluzioni deboli (o distribuzionali)" dei sistemi differenzialparziali lineari nasce da circostanze di questo genere; ed è noto che tali soluzioni possono sempre essere approssimate da funzioni regolari con precisione arbitraria. È anche ovvio che se una modifica in tal senso non viene in qualche modo incorporata negli algoritmi che si disegnano per risolvere un problema equazionale che non ha soluzioni, quegli algoritmi sono automaticamente infondati, e che le risposte che se ne derivano possono essere parzialmente o totalmente illusorie.

Non resta che rinviare il lettore curioso di approfondimenti alla Bibliografia Generale (anch'essa piuttosto esemplificativa) per quanto riguarda le intere classi di problemi funzionali-numerici di cui non abbiamo nemmeno accennato, nonché la effettiva costruzione di algoritmi di soluzione delle (1) e/o (2) nelle loro molte particolari realizzazioni.

BIBLIOGRAFIA GENERALE

BALAKRISHAN, A.V.: “Applied Functional Analysis”, Springer 1976

BERGER, M.S.: “Non linearity and functional Analysis”, Acad. Pr. 1977

COLLATZ, L.: “Functional Analysis and Numerical Mathematics”, Acad. Pr. 1966

DAUTRAY, R., LIONS, J-L : “Mathematical Analysis and Numerical Methods” (engl. transl.) 6 vols, Springer 1988-1990 (édit. franç. 1984-85)

DIEUDONNÉ, J.: “Foundations of Modern Analysis”, Acad. Pr. 1969

KANTAROVICH, L.V., AKILOV, G.P.: “Functional Analysis in Normed Spaces”, Pergamon Pr. 1964

KRASNOSELSKII, M.A. et al.: “Approximate Solutions of Operator Equations” (engl. transl.) Noordhoff 1972

MARCHUK, G.I.: “Methods of Numerical Mathematics”, (engl. transl.) Springer 1982

TRENOGUINE, V.: “Analyse fonctionnelle” (trad. franç.) MIR 1985

VARGA, R.: “Functional Analysis and Approximation Theory in Numerical Analysis”, The Soc. for Ind. and Appl. Maths, 1971

ZEIDLER, E.: “Nonlinear Functional Analysis”, 5 vols, Springer 1986