Momenti di svolta nello sviluppo del pensiero matematico

La scoperta del metodo assiomatico: Euclide, Hilbert, Gödel

(lezioni di Modena, ottobre-dicembre 2006)

Gabriele Lolli

Dipartimento di Matematica, Torino

Programma dei sei incontri

1. Il metodo ipotetico-deduttivo di fine Ottocento
2. Da Euclide a Hilbert
3. Il problema della completezza
4. La nozione di "formale"
5. Incompletezza e indecidibilità
6. La metamatematica

Indice

I temi trattati in questi incontri sono quelli del volume *Da Euclide a Gödel*[1], presentati per ragioni di tempo e di forma in un ordine e con enfasi differente.

Se in queste lezioni parliamo dell'assiomatica, non è per suggerire che questo metodo riassuma o coinvolga tutti gli aspetti del pensiero matematico. La scelta dipende dalle competenze di chi parla, anche se ci sono validi motivi per giudicare che il metodo assiomatico, in particolare dopo la svolta di fine Ottocento, sia un elemento fondamentale per capire la natura della matematica. Ma non bisogna pensare che sia l'unico; in particolare proprio a fine Ottocento la codifica moderna del metodo assiomatico è stata accompagnata da un'altra concezione, il riduzionismo collegato alla teoria degli insiemi, che spinge in un'altra direzione (allora si contrapponeva il metodo genetico al metodo assiomatico) e ancora è considerato un'alternativa; la conciliazione pratica che si è realizzata attraverso il silenzio è fortemente insoddisfacente, a uno spirito critico.

La riflessione che segue dunque non deve essere interpretata come un invito a pensare la matematica come un insieme di sistemi assiomatici, tanto meno nell'insegnamento.

Tuttavia il fatto che si debba spaziare su più di duemila anni di storia per capire in cosa consista il metodo assiomatico dovrebbe di per sé suggerire che si tratta di una conquista sofisticata e faticosa, e che merita di essere conosciuta bene, almeno da parte di chi si occupa di matematica. Quello che si intende normalmente quando si parla dell'importanza di vedere una situazione in termini assiomatici è solo l'attitudine a isolare bene le assunzioni rilevanti. Lo storico E. T. Bell diceva che quello che aveva imparato da Euclide era che senza assunzioni non c'è dimostrazione - un po' poco[2]

La parla "svolta" potrebbe essere sostituita da altre, come "rivoluzione", o "frattura" o "progresso": ma non vogliamo entrare in una discussione di filosofia della storia. Piuttosto dobbiamo imparare a capire perché in determinati momenti si senta la necessità da parte dei matematici di elaborare

[1] G. Lolli, *Da Euclide a Gödel*, Il Mulino, Bologna, 2004.

[2] Anche un ottimo matematico e scrittore di matematica come Ian Stewart, nel suo recente *Com'è bella la matematica*, Bollati Boringhieri, Torino, 2006, quando parla degli assiomi dice che gli assiomi si Euclide si possono accettare o no, se si crede: "Euclide vuole semplicemente rendere esplicite le regole del suo gioco" (pag. 76). Euclide non aveva letto Wittgenstein. Valenti matematici quando parlano di questioni fondazionali si esprimono in modo così *casual*, perché ritengono di poterle capire senza studiare; l'impressione che danno è come minimo quella di un difetto di stile e di eleganza.

considerazioni metodologiche, invece che, o oltre a fare matematica, e in risposta a quali problemi quelle considerazioni abbiano prodotto quello hanno prodotto.

Partiremo dalla descrizione del metodo assiomatico come viene codificato nel cosiddetto metodo ipotetico-deduttivo di fine Ottocento. Esso rappresenta in effetti una svolta, ed è a questa che si riferisce Dieudonné quando afferma che "ad ogni matematico che abbia a cuore la probità intellettuale s'impone ormai la necessità assoluta di presentare i propri ragionamenti in forma assiomatica ... con parole che si sono svuotate di ogni significato intuitivo"[3].

Si presenterà quindi la giusta curiosità di sapere cosa e come facevano prima i matematici, poi anche di cosa e come hanno fatto dopo.

Si può anticipare che guardando al passato possono sorgere perplessità sulla dichiarazione di svolta, perché la presentazione della matematica occidentale è sempre stata legata a un cosiddetto metodo assiomatico risalente ad Euclide. Esso tuttavia era diverso nella sostanza, se non nella terminologia, da quello contemporaneo. Ma doveva essere causa di una coscienza infelice[4] perché tale presentazione, o immagine, o meglio l'ideale e il modello di riferimento assiomatico non corrispondeva per nulla alla pratica, alla reale dinamica della scoperta e dello studio di nuova matematica.

I vari punti che si devono chiarire allora, per avere una visione corretta e coerente, anche se non completa, dello sviluppo della matematica, sono, almeno, i seguenti:

1. Il metodo assiomatico di Euclide, e la sua novità rispetto a una matematica che aveva comunque dietro le spalle alcuni secoli di sviluppo.

2. Quello che restava fuori, già allora, dal quadro euclideo, e quello che entrando in seguito a far parte della matematica è cresciuto ed è stato riconosciuto valido e corretto, e di fatto accettato come matematica, anche al di fuori dell'impianto assiomatico.

3. Cosa è che manda in crisi una pratica comunque feconda e che era stata capace fino ad allora di ignorare o scantonare dalle lacune della sua organizzazione.

[3] J. Dieudonné, "Les méthodes axiomatiques modernes et les fondements des mathématiques", in *Les grands courants de la pensée mathématique*, a cura di F. Le Lionnais (1948), seconda edizione arricchita Paris, A. Blanchard, 1962, pp. 543-55, cit. p. 544.

[4] In senso hegeliano, cioè per lo Spirito, non per i singoli.

4. Perché questa crisi porta proprio alla particolare riformulazione del metodo assiomatico che ci ritroviamo, in un modo che non è tanto diverso dal punto di vista esteriore, ma con sostanziali differenze di concezione, e quali.

5. Cosa si può dire sulle aspettative dei matematici di fine Ottocento e sulla matematica in genere quando con gli strumenti logici si può incominciare a fare uno studio metamatematico delle teorie assiomaticamente organizzate. Risulterà infatti che il ripensamento dell'organizzazione assiomatica rende più evidenti i legami con la logica formale, e che questa nello stesso periodo si attrezza per affrontare lo studio delle teorie assiomatiche.

Come si evince da quanto premesso, non trattiamo questi temi nell'ordine temporale. Iniziamo dalla svolta di fine Ottocento perché è quella che si fa sentire ancora oggi, e perché la si può seguire bene, nel contesto dello sviluppo della matematica, sulla base solo delle nostre conoscenze scolastiche. Di solito si parte da Euclide, ma Euclide non è il *big bang*: la sua svolta rispetto alla matematica precedente, greca e orientale, è stata forse ancora più rivoluzionaria di quella moderna, ma è un mistero: a parte che è difficile da analizzare se non si conosce la matematica pre-euclidea, a partire dai Pitagorici, e senza prendere in esame anche il contesto filosofico, resta comunque in definitiva inspiegata.

Accenneremo anche brevemente alla alternativa del metodo genetico, e alle critiche di coloro che rifiutano l'assiomatica come sistemazione definitiva della matematica.

Come riferimenti bibliografici indichiamo per la storia moderna e contemporanea i volumi di U. Bottazzini, *Il flauto di Hilbert*, Utet Libreria, Torino, 1990 e J. Dieudonné (a cura di), *Abrégé d'histoire des mathématiques*, 2 voll., Hermann, Paris, 1978.

1 Il metodo ipotetico-deduttivo

Il metodo assiomatico a proposito del quale Dieudonné afferma che si impone *ormai*[5] con necessità assoluta è piuttosto recente, per quanto l'informazione possa suonare sorprendente, a chi risulta nuova: risale a poco più di cento anni fa.

Presentiamo la nascita e la formulazione di questa impostazione con le parole di Federigo Enriques, non tanto per ricorrere ad un argomento *ad auctoritatem*, quanto perché ci sembra la presentazione più chiara che sia stata data, e da un matematico attivo, che colloca questo sviluppo nelle esigenze della crescita della matematica del secolo decimonono.

Enriques svolge le sue considerazioni nel volume *Per la storia della logica*, del 1922[6]; con "logica" Enriques non intende la logica formale, ma la teoria della scienza, o la metodologia, in particolare delle scienze dimostrative, cioè della matematica.

Afferma Enriques che lo sviluppo della logica anteriore al secolo decimonono non ha mutato, apparentemente, il concetto tradizionale dell'ordinamento delle scienze dimostrative, ma "ora [seconda metà dell'Ottocento] accade che la riforma della logica, lungamente preparata, maturi qui, sotto l'impulso di diversi motivi, attinenti allo sviluppo delle Matematiche".

Le spinte che hanno portato a quella che Enriques chiama "riforma della logica", o anche "dell'ordinamento della scienza deduttiva", sono "alcuni movimenti di pensiero, di origini in gran parte distinte che pure interagiscono e si incontrano, in un medesimo concetto riformatore".

Enriques indica:

1. il nascere della geometria proiettiva;

2. le geometrie non euclidee;

3. l'algebra simbolica inglese;

4. l'interpretazione positivista delle teorie fisiche come modelli;

5. il lavoro volto a "dare solido fondamento all'Analisi, superando le difficoltà ormai mature del Calcolo infinitesimale e sciogliendo i paradossi

[5]C.vo nostro.

[6]F. Enriques, *Per la storia della logica*, Zanichelli, Bologna, 1922, ristampa anastatica 1987. Parafrasiamo e citiamo dal Cap. III, dal paragrafo 20 in avanti.

offerti dalle serie divergenti, le pseudo-dimostrazioni dei massimi e minimi, delle derivate ecc.", vale a dire il processo noto come *aritmetizzazione* dell'Analisi e, "in margine all'accennato movimento", l'analisi dell'infinito, con il che Enriques intende la teoria degli insiemi di Cantor e Dedekind;

6. ma "la riforma della logica contemporanea si afferma pienamente soltanto attraverso la *critica più recente dei principi della geometria*, per la quale i pensatori matematici acquistano coscienza matura del significato di una rivoluzione compiuta nei secoli".

Non tutti sono d'accordo con quest'ultimo giudizio di Enriques, per alcuni l'algebra astratta, di cui parleremo, è stata un fenomeno più decisivo, ma Enriques era un geometra, e gli si può concedere un pregiudizio settoriale. Comunque il suo giudizio è politicamente efficace, e accettabile, se non del tutto storicamente corretto: è importante che egli radichi in modo convincente la rivoluzione nel filone di quella che si chiama *mainstream* della matematica.

L'algebra, quella di cui parleremo, non quella che era inglobata nell'Analisi, era allora marginale - come lo è anche talvolta la sua discendente, algebra astratta, in giudizi recenti - al punto che chi la coltivava doveva impegnarsi a confutare la "tendenza . . . a rifiutare il punto di vista che considera l'Algebra come una scienza, *in qualche senso analoga alla geometria*, e [la tendenza] ad adottare una o l'altra di due diverse concezioni, che considerano l'Algebra come un'*Arte*, oppure come un *Linguaggio*: come un Sistema di Regole, oppure come un Sistema di Espressioni, ma non come un Sistema di *Verità*[7].

Passiamo brevemente in rassegna, sia con le parole di Enriques, sia con qualche integrazione, ciascuna di queste novità della matematica dell'Ottocento. L'elenco non esaurisce tutte quelle rilevanti, ad esempio si potrebbe accennare alla storia della teoria degli invarianti, ma si tratta senza dubbio delle più importanti, e già così il loro intreccio è sufficientemente complicato.

[7]W. R. Hamilton, "Theory of conjugate functions or algebraic couples, with a preliminary and elementary essay on algebra as the science of pure time", *Trans. Royal Irish Acad.*, 17, 1837, pp. 293 - 422; passi scelti riprodotti in W. Ewald (ed.), *From Kant to Hilbert*, Clarendon Press, Oxford, 1996, vol. I, pp. 362-441.

1.1 La geometria proiettiva

La nuova disciplina cresce a partire dall'inizio dell'Ottocento, con Monge e Poncelet, dopo qualche anticipazione in Desargues e Pascal. Enriques dà grande importanza al principio di dualità di Gergonne (1819).

Gergonne osserva che i teoremi della geometria di situazione (non metriche) si presentano a coppie e si scambiano uno nell'altro scambiando punti con rette nel piano o punti con piani nello spazio, rette invariate.

Già altri prima di lui avevano notato la possibilità di ottenere un risultato da un altro, ad esempio Brianchon (1806) aveva dedotto il teorema sull'esagono circoscritto ad una conica da quello di Pascal sull'esagono inscritto[8], con una trasformazione polare della figura. Gergonne fu anzi accusato di plagio per essersi attribuito il merito dell'osservazione della dualità. Poncelet sosteneva che la spiegazione della dualità stava nella relazione polo-polare da lui messa al centro delle sue indagini.

Gergonne osserva anche che "esiste per di più tra le dimostrazioni di una stessa coppia la stessa corrispondenza che tra i loro enunciati". Secondo Enriques Gergonne mostra tuttavia solo lo sviluppo parallelo dei primi teoremi, ma non spinge l'analisi a mostrare che si ha un sistema di postulati sufficiente a edificare la geometria di situazione.

Il ragionamento di Gergonne presuppone lo sviluppo successivo della geometria proiettiva a opera di Staudt (1847), che la rende indipendente da nozioni metriche ancora mescolate fino ad allora con quelle di situazione. Ancora più importante è stata l'introduzione delle coordinate di Plücker.

"Plücker fa riposare la legge di dualità sulla considerazione delle coordinate di rette e di piani, che permette un identico trattamento analitico delle relazioni correlative: le proprietà analitiche delle terne di numeri (x, y, z) si rispecchieranno, sia nelle figure dello spazio dove codeste terne vengano prese come coordinate dei loro elementi generatori 'i punti', sia nei sistemi di cerchi del piano, qualora gli stessi numeri vengano assunti come coefficienti dell'equazione di un cerchio, ossia come coordinate di cerchi ecc."

Dagli elementi di proiettiva è noto che nell'equazione $u_1x_1+u_2x_2+u_3x_3 = 0$ di una retta del piano la completa simmetria di u_i e x_i permette di considerare le u_i come variabili e le x_i come coefficienti, e l'equazione come rappresentante il fascio di rette per il punto $\langle x_1, x_2, x_3\rangle$.

[8]Il teorema di Pascal afferma che dato un esagono inscritto in una conica, i tre punti di intersezione delle coppie di lati opposti sono allineati; il teorema di Brianchon, per un esagono circoscritto, afferma che le tre diagonali principali si incontrano in un punto.

A Gergonne si deve anche il primo uso, e menzione, della nozione di "definizione implicita", che arricchisce il dibattito sulla natura delle definizioni, ferma alla distinzione tra definizioni reali e definizioni nominali della tradizione filosofica, con pochi arricchimenti. La nozione servirà a spiegare, non senza contrasti, il tipo di definizione offerta dai sistemi di postulati, se di definizione si tratta, in sostituzione delle definizioni reali della pratica precedente.

Secondo Enriques "la teoria della definizione implicita d'un sistema di concetti per mezzo di un sistema di proposizioni, è divenuta essenziale per la logica contemporanea. Ma essa non avrebbe potuto apparire nella luce in cui oggi la vediamo, se non risultasse chiarita da quel principio generale di sostituibilità dei concetti, che ha il suo germe nel principio di dualità della geometria proiettiva". Il principio di sostituibilità dei concetti, come vedremo, è ciò che rende possibili diverse interpretazioni per uno stesso sistema di assiomi.

1.2 Le geometrie non euclidee

Va da sé che le geometrie non euclidee, che rappresentano "una *possibilità* geometrica che non si accorda con la nostra intuizione dello spazio", hanno reso impossibile continuare a pensare che i concetti della geometria fossero legati allo spazio. Per Enriques è un colpo mortale per il razionalismo, in quanto sono la prova che la realtà non può essere determinata *a priori* ma solo con una verifica sperimentale. I concetti geometrici euclidei sono eventualmente legati solo alla nostra intuizione dello spazio; si tratta di un'intuizione di esseri limitati e condizionati dalla loro natura fisica e biologica e dalla loro situazione particolare nel mondo.

Enriques non si sofferma molto su questo sviluppo, se non per ricordare la ricerca di possibili interpretazioni concrete per le geometrie non euclidee, suggerita da Riemann e messa in atto da Beltrami[9], ad esempio con la pseudosfera[10]: "da questo lavoro ... si svolge, in tutta la sua pienezza, il concetto

[9] E. Beltrami, "Saggio d'interpretazione della geometria non-euclidea", *Giornale di matematiche ad uso degli studenti*, vol. 6, 1868, pp. 284-312.

[10] Ricordiamo che Beltrami ottenne con superfici a curvatura costante negativa modelli di una porzione del piano iperbolico di Lobachevski. Klein nel 1871 descrisse il modello nel quale i punti sono i punti interni a una circonferenza e le rette sono le corde. Poincaré nel 1882 descrisse due modelli: nel primo i punti sono i punti interni a una circonferenza e le rette sono gli archi di cerchio perpendicolari al contorno e i diametri; nel secondo i

della *geometria astratta*, in cui si può ravvisare il naturale prolungamento del principio di dualità guadagnato nella geometria proiettiva, ed anche il preludio immediato della nuova concezione del *sistema ipotetico-deduttivo*, accolta nella logica contemporanea".

1.3 L'algebra simbolica

Nel corso dell'Ottocento sono elaborate, soprattutto in Inghilterra, nuove teorie algebriche, il cui oggetto non sono i numeri.

Per Enriques si tratta di "una pluralità di indirizzi apparentemente senza un criterio direttivo, con le Matematiche che affermano il diritto ad esistere come dottrina pura, indipendente dalle applicazioni alla scienza naturale".

Qualche criterio tuttavia si distingue; le nuove teorie non sono frutto di decisioni arbitrarie - se mai in matematica lo sono - ma rispondono a esigenze interne alla crescita della matematica. Particolarmente influente si rivela una serie di osservazioni, espresse da varie parti, sul calcolo delle operazioni.

Si notavano ad esempio analogie tra potenze, derivate e altre operazioni, come l'iterazione di una funzione, del tipo

$$x^n \cdot x^m = x^{m+n}$$

$$f^n(f^m(x) = f^{m+n}(x)$$

$$\tfrac{d^n}{dx^n}\left(\tfrac{d^m y}{dx^m}\right) = \tfrac{d^{n+m}y}{dx^{n+m}}$$

Nel concentrarsi sullo studio di queste analogie, si parlava di un "principio di separazione dei simboli di operazione da quelli di quantità"[11], e si vedeva l'algebra sempre più orientata ai primi.

Dal principio di separazione dei simboli di operazione molti "furono condotti all'idea che non era la natura degli oggetti in esame che era la più significativa, ma piuttosto le leggi di combinazione dei loro simboli"[12].

Per questo salto non erano ovviamente sufficienti le osservazioni di analogie formali, ma agivano altri stimoli. Fino ad allora la matematica era

---

punti sono i punti di un semipiano e le rette sono o gli archi di cerchio ortogonali alla retta che delimita il semipiano e le rette perpendicolari. La geometria ellittica di Riemann ha un modello sulla superficie sferica.

[11] D. Gregory, "On the real nature of symbolical algebra", *Trans. Royal Society Edimburgh*, vol. 6, 1840, pp. 208-16, parzialmente riprodotto in Ewald, cit.

[12] E. Koppelman, "The calculus of operations and the rise of abstract algebra", *Archive of History of Exact Sciences*, 8, 1972, pp. 155-242.

ancora definita come la scienza delle grandezze, o delle quantità. Ci sono varie specie di quantità, tutto ciò che è passibile di crescere o di diminuire, così tante che non è possibile enumerare tutte, "e questa è l'origine dei vari rami della matematica, ciascuno interessato a un particolare genere di grandezza" diceva Eulero[13]. Anche Gauss era dell'idea che "la matematica insegna in effetti verità generali sulle relazioni tra quantità"[14]; benché affermasse che "i matematici fanno completamente astrazione dalla natura degli oggetti e dal significato delle loro relazioni; non c'è che da enumerare queste relazioni e confrontarle tra loro"[15], Gauss intendeva forse che nella matematica restasse solo l'aspetto numerico o geometrico delle relazioni tra gli oggetti.

A scuotere queste credenze tramandate contribuiscono diversi fattori. Non è da trascurare il lavoro fatto sui tradizionali sistemi di algebra; Lagrange si era accorto che la regola dei segni nella estensione delle operazioni agli interi era una conseguenza della legge distributiva. Proprio i numeri negativi erano una spina nella ripetuta concezione delle quantità. La sistemazione dei numeri negativi e di quelli immaginari proponeva il problema delle regole da assumere nel calcolo. La presentazione di Lagrange del calcolo infinitesimale era essenzialmente algebrica, ancor più di quanto potessero suggerire le leggi delle derivate.

In Inghilterra all'inizio dell'Ottocento si forma un movimento per svecchiare la matematica legata a Newton e chiusa agli sviluppi del continente, dove con le notazioni di Leibniz prevaleva l'aspetto algebrico su quello geometrico delle flussioni; ma su molte questioni il modo di procedere dei continentali era per altri versi approssimativo e insoddisfacente, addirittura a proposito dei numeri.

Un personaggio come Cauchy si esprimeva ancora dicendo che "le grandezze che servono per aumentare sono indicate da numeri preceduti dal segno $+$, e le grandezze che servono come diminuzione sono indicate da numeri preceduti dal segno $-$ ... e i simboli $+$ e $-$ messi davanti a numeri si concepiscono come aggettivi posti davanti ai nomi"[16]. Si usavano locuzioni inconsistenti

[13] L. Euler, *Vollständige Anleitung zur nieder und höheren Algebra*, 1770, trad. inglese *Elements of Algebra*, Springer, New York, 1984, pag. 1.

[14] C. F. Gauss, "Fragen zur Metaphysik der Mathematik", *Werke*, X/1, Gottingen, 1870-1933, pp. 396-7, parzialmente riprodotto in Ewald, cit.

[15] C. F. Gauss, *Werke*, II, pag. 176.

[16] G. Peacock, *Report on the recent progress and present state of certain branches of analysis*, The Third Meeting of the British Association for the Advancement of Science, Cambridge 1833, London 1834, pp. 185-352, cit. p. 193.

come “sottrarre una grandezza maggiore da una minore” per spiegare la generazione dei numeri negativi. Si collocavano i negativi dopo $+\infty$, in base al principio che siccome dividendo un numero per numeri positivi sempre più piccoli si avevano risultati sempre più grandi, quando si passava a numeri ancora minori, i negativi, minori di 0, il risultato doveva essere maggiore dell'infinito. Le incertezze sui numeri immaginari si riflettevano su rinnovati dubbi sui numeri negativi (le “false” soluzioni).

Dei numeri immaginari si avevano solo interpretazioni geometriche, a opera di Argand (1806) e Gauss (1831). In alternativa altri, come F.-J. Servois (1767-1847) affermavano che l'unica legittimazione degli immaginari era la loro utilità nei calcoli, e quindi potevano restare e restavano senza interpretazione.

All'interno del movimento per la riforma dell'Analisi, la “Analytical Society”, George Peacock si dedica a correggere la tradizionale avversione inglese per l'algebra con una trattazione sistematica dei suoi principi. Nel 1833 propone una distinzione tra algebra aritmetica e algebra simbolica. Nell'algebra aritmetica le variabili rappresentano numeri reali positivi e le operazioni si assumono nel senso ordinario. Nell'algebra simbolica si elimina ogni restrizione, sia sulle variabili sia sulle operazioni (ad esempio che la sottrazione si possa eseguire solo se il minuendo è maggiore del sottraendo). Le regole a cui sono soggette le operazioni sono libere dalle restrizioni delle loro definizioni, definizioni che alcune volte mancavano o erano solo geometriche.

Nell'algebra simbolica le operazioni si potevano definire in base al *popular meaning*, ma questo non rappresentava né una deduzione delle stesse né una loro fondazione, quanto solo un suggerimento; poi erano soggette solo alle condizioni simboliche: “Il significato delle operazioni eseguite, così come i risultati ottenuti ... deve essere derivato non dalle loro definizioni o dai significati assunti” ma dalle regole postulate[17]. I risultati ottenuti potevano essere applicati all'algebra aritmetica attraverso una interpretazione. Ma le loro manipolazioni erano più estese, e magari ne venivano conclusioni che nell'interpretazione non erano possibili.

Qualche volta Peacock accenna alla possibilità di un sistema di algebra simbolica che non abbia alcuna interpretazione, ma benché lo dichiari ammissibile lo ritiene di scarso interesse, e considera l'algebra aritmetica una fonte di ispirazione. Quando l'interpretazione sussiste, deve esserci perfetta consonanza tra algebra simbolica e aritmetica, requisito che Peacock chia-

[17]G. Peacock, *Report on the recent progress ...*, cit.

ma "principio di permanenza di forme equivalenti", e che gli impedisce di prendere in considerazione vere deviazioni dall'algebra aritmetica.

Più coraggioso, il suo allievo Gregory si spinge oltre. Gregory usa davvero nelle espressioni le lettere non solo per i numeri ma per operazioni qualunque, ponendole nelle posizioni dei numeri. Egli scrive ad esempio

$$f(x+h) = e^{h\frac{d}{dx}} f(x)$$

grazie al fatto che

$$e^x = 1 + x + \tfrac{x^2}{2!} + \ldots.$$

In modo coerente, nel considerare leggi come $a^m \cdot a^n = a^{m+n}$ e $(a^m)^n = a^{mn}$ Gregory passa a discutere quale senso di può dare a $(+)^m$ e $(-)^m$ quando m sia non solo un numero razionale qualunque, ma anche a sua volta un simbolo di operazione, come $(+)^{log}$, pur ammettendo di non avere ancora la risposta.

Tra gli esempi dell'effetto del principio di separazione, Gregory propone

$$(1.) FF(a) = F(a) \quad (2.) ff(a) = F(a)$$

$$(3.) Ff(a) = f(a) \quad (4.) fF(a) = f(a)$$

e osserva che tra le operazioni aritmetiche, due soddisfano le leggi elencate, quelle indicate dai simbolo $+$ e $-$; ma esistono anche due operazioni geometriche soggette alle stesse leggi: Gregory si esprime in modo fantasioso, come "il girare di un punto per una circonferenza", ma sembra di capire che si tratti di una rotazione rispettivamente di 360° e 180° gradi. "Non esiste alcuna analogia reale tra le due operazioni aritmetiche e le due operazioni geometriche . . . in quanto le due operazioni non si possono neanche considerare l'una l'opposta dell'altra [in geometria]. La relazione che esiste è dovuta non a una identità della loro natura, ma al fatto di essere combinate tra loro dalle stesse leggi".

Gregory ripete da Peacock che "l'algebra simbolica è la scienza che tratta la combinazione di operazioni definite non dalla loro natura [. . .] ma dalla leggi di combinazione alla quale sono assoggettate"[18].

Le definizioni sono diverse dalle regole. "Un simbolo è *definito* quando sono esplicitate per il suo uso regole che ci permettano di accettare o rifiutare qualsiasi proposta trasformazione che lo coinvolge. Un simbolo semplice è

[18]Gregory, cit.

spiegato quando ad esso è assegnato un significato che ci permetta di accettare o rifiutare l'applicazione della sua definizione, come conseguenza di quel significato"[19].

In questa prospettiva si moltiplicano le algebre.

William Hamilton propone nel 1837 un'algebra di coppie per i complessi, ma basata su un sistema simbolico di enti che interpreta come istanti temporali[20]. Boole costruisce nel 1847 la sua algebra per le leggi del pensiero.

Ogni algebra è una scienza dei simboli costruita con le sue regole, e poi applicata con un'interpretazione.

Con l'affermarsi dell'autonomia degli apparati simbolici diventa scottante il problema del significato. Ma i protagonisti non hanno una lucida consapevolezza metodologica immediata e non lanciano un chiaro programma; sono vincolati dalla saggezza ricevuta.

"Nessun algebrista della scuola di Cambridge sembra ammettere l'idea di una struttura algebrica che non sia concepita in rapporto a un'interpretazione privilegiata"[21].

Hamilton ricorda di essersi interessato alla dottrina dei numeri negativi e immaginari perché, d'accordo con quelli che ritenevano che non si potessero considerare quantità, era insoddisfatto di ogni spiegazione che non desse loro, dall'inizio, una chiara interpretazione un significato.

Hamilton arriva addirittura a insistere, pur nel contesto simbolico, che l'algebra abbia un suo oggetto che la caratterizza come scienza.

"Si potrebbe provare un naturale rimpianto, [per] il destino dell'Algebra, se uno studio che impegna sempre di più i matematici, al punto da aver quasi sopravanzato lo Studio della Scienza Geometrica, si trovasse alla fine a non essere per nulla lo Studio di una Scienza, in ogni senso stretto e proprio ... Spera perciò di incontrare comprensione chiunque voglia indagare se l'Algebra esistente, nello stato in cui è stata sviluppata dai maestri delle sue regole e del suo linguaggio, non offra invero alcun rudimento che possa incoraggiare una speranza di sviluppare una Scienza dell'Algebra: una

[19]A. De Morgan, "On the foundation of algebra", *Trans. Cambridge Philosophical Society*, vol. 7, 1842, pp. 173-87, parzialmente riprodotto in Ewald, cit.

[20]W. R. Hamilton,"Theory of conjugate functions or algebraic couples" cit. $A < B$ significa che l'istante A precede l'istante B, la coppia $\langle A, B\rangle$ denota un intervallo: Hamilton costruisce così i numeri naturali, che chiama "ordinali".

[21]M. Mugnai, "Alle origini dell'algebra della logica", in V. M. Abrusci, E. Casari, M. Mugnai (a cura di), *Atti del Convegno Internazionale di Storia della Logica*, CLUEB, Bologna, 1983, pp. 117 - 31.

scienza propriamente così chiamata; rigorosa, pura e indipendente; dedotta con ragionamenti validi dai suoi principi intuitivi; e perciò oggetto di una contemplazione a priori non meno della Geometria, né meno distinta, nella sua essenza, dalle Regole che può insegnare o usare, e dai Segni con i quali può esprimere il suo significato. L'Autore di questo saggio è pervenuto alla convinzione che tale rudimento consista nella Intuizione del Tempo"[22]. L'oggetto dell'algebra per Hamilton non è la quantità, quanto l'Ordine in Progressione.

La pretesa di trovare un oggetto per l'Algebra si scontra tuttavia con la moltiplicazione dei sistemi algebrici devianti rispetto a quello dell'algebra aritmetica.

Si capisce l'importanza che in questo clima ha avuto la scoperta dei quaternioni da parte di Hamilton stesso nel 1843: un sistema di enti $a+ib+jc+kd$ con

$$i^2 = j^2 = k^2 = -1$$
$$ij = -ji = k, \quad jk = -kj = i, \quad ki = -ik = j.$$

Il sistema dei quaternioni smentiva l'idea che un sistema di numeri debba necessariamente comportare una moltiplicazione commutativa. Hamilton voleva definire attraverso di essi le rotazioni nello spazio, come i complessi forniscono le rotazioni nel piano. I quaternioni ebbero una certa importanza nella nascita dell'analisi vettoriale, anche se meno delle speranze del suo inventore.

Nello stesso tempo, o poco dopo, si scopriranno vari tipi di prodotto tra vettori, tra cui il prodotto esterno, con Grassman (1844), le matrici con un'altra moltiplicazione non commutativa[23], i divisori dello zero nei biquaternioni[24], una moltiplicazione neanche associativa negli ottetti, o ottave, di Grave e Cayley, coppie di quaternioni, con una moltiplicazione $\langle q_1, q_2\rangle\langle q_3, q_4\rangle = \langle q_1q_3 - \overline{q}_4q_2, \; q_2\overline{q}_3 + q_4q_1\rangle$, dove $\overline{q}$ è il coniugato di q.

Verranno poi le algebre di Clifford, sistemi dove si parte da n unità $1, e_1, \ldots, e_{n-1}$ con $e_i^2 = -1$ e $e_ie_j = -e_je_i$ e per ognuno dei 2^n sottoinsiemi di n di pone $e_H = e_{i_1}e_{i_2}\ldots e_{i_h}$ con $i_1 < \ldots i_h$, ottenendo una moltiplicazione associativa.

[22] W. R. Hamilton,"Theory of conjugate functions or algebraic couples" cit.

[23] A proposito delle matrici, occorrerebbe ricordare l'importante sviluppo matematico della teoria degli invarianti, con i suoi legami con la geometria

[24] Sono i quaternioni definiti sul corpo base $\mathbb{C}$ invece che $\mathbb{R}$; il sistema equivale a quello delle matrici di ordine 2 su $\mathbb{C}$, come osservato da Cayley e Peirce.

Questi sistemi non nascono tutti per decisioni volontaristiche, ma nella ricerca di strumenti per la soluzione di problemi matematici.

Fino a metà del secolo, gli unici gruppi finiti che si considerano sono quelli di permutazione, in connessione alle ricerche sulla soluzione delle equazioni algebriche; nel 1854 Cayley darà la prima definizione moderna di gruppo astratto finito[25].

Non ha più tanto senso continuare a usare i vecchi simboli $+$ e $\cdot$. Nel 1867 Hankel introdurrà la nozione astratta di legge di composizione[26]. Un'anticipazione di questo termine era apparsa nel 1822 a opera, anonima, di Gergonne.

Come nell'esperienza della geometria, ma in modo più marcato, nel campo dell'algebra si ha un duplice fenomeno. Da una parte compaiono diversi sistemi algebrici, dall'altra ciascuno di essi è passibile di tante interpretazioni.

In geometria, le interpretazioni sono inizialmente vissute secondo Enriques come un cambiamento di senso, un "leggere per", come dirà anche Beltrami.

Nell'algebra simbolica, ogni sistema è autonomo, e inizialmente ha la sua interpretazione privilegiata, ma "ogni interpretazione che non infici la verità delle relazioni supposte è ugualmente ammissibile"[27].

Alla fine del secolo, le interpretazioni diventeranno "sistemi di enti", o "insiemi", sui quali sono introdotte operazioni soddisfacenti determinate regole.

Gregory appare consapevole dei vantaggi del trasferimento di risultati da una interpretazione ad un'altra.

"Queste leggi sono state in molti casi suggerite ... dalle leggi delle note operazioni numeriche; ma il passo che è compiuto nel passaggio dall'algebra aritmetica all'algebra simbolica consiste in questo che, tralasciando la natura delle operazioni che i simboli usati rappresentano, noi supponiamo l'esistenza di classi di operazioni sconosciute soggette alle stesse leggi. Siamo allora in grado di provare certe relazioni tra le differenti classi di operazioni, che, quando sono espresse tra i simboli, sono chiamate teoremi algebrici. E se possiamo mostrare che certe operazioni di qualsiasi genere in una scienza sono soggette alle stesse leggi di combinazione di queste classi, i teoremi sono veri di esse, come incluse nel caso generale"[28].

[25] Per la definizione generale si deve aspettare il trattato di H. Weber del 1899.

[26] H. Hankel, *Théorie des systèmes de nombres complexes*, Voss, Leipzig, 1867.

[27] G. Boole, *The Mathematical Analysis of Logic*, London - Cambridge 1847, p. 3.

[28] Gregory, cit.

De Morgan ha introdotto, come una battuta, la locuzione di "simboli svuotati di significato". Ma calcoli non interpretati significa per adesso solo che non ci si pensa.

Si intersecano allora due aspetti, quello del carattere formale delle regole o delle assunzioni, e quello del carattere formale del ragionamento, su cui torneremo; dopo un laborioso parto, si arriverà a riconoscere che in matematica: "Il ragionamento è formale nel senso che il significato delle proposizioni non entra assolutamente in questione"[29].

Questo è un filone che si potrebbe risalire all'indietro attraverso Condillac fino a Leibniz e forse più indietro ancora. Leibniz aveva riconosciuto la subordinazione "dell'algebra speciosa alla speciosa generale, della scienza delle formule significanti la quantità alla dottrina delle formule o espressioni in generale dell'ordine, della similitudine, della relazione"[30].

Condillac, in *La langue des calculs* aveva detto come quando si deve risolvere un'equazione come $x + a - b = c$ non abbiamo bisogno di sapere cosa indicano le lettere da cui è formata, e se lo sapessimo non ci penseremmo, e solo dopo mettiamo dei valori, e le operazioni per questo sono fatte in modo meccanico.[31].

Non sorprende quindi di ritrovarlo anche in campo geometrico. Poncelet si proponeva "di ridurre, in qualche modo, a un puro meccanismo, a una semplice sostituzione di nomi e lettere, scritti al posto l'uno dell'altro, la traduzione di tutte le proprietà ... che appartengono a una figura a alla sua reciproca" con trasformazione polare[32].

Gergonne teorizza in generale: "Si ripete in continuazione *che non bisogna ragionare su oggetti dei quali non sia abbia un'idea precisa*; e invece, nulla vi è di più falso. Si ragiona in effetti con le parole proprio come in algebra si calcola con le lettere; e come si può eseguire con precisione un calcolo ... senza interrogarsi ... sul significato dei simboli con i quali si opera, si può ... seguire un ragionamento senza conoscere ... il significato dei termini nei quali si esprime, o senza farvi caso se lo si conosce"[33].

---

[29]A. N. Whitehead, *A Treatise on Universal Algebra, with Applications*, Cambridge Univ. Press, Cambridge, 1898, p. vi.

[30]W. G. Leibniz, *Mathematische Schriften*, Gerhardt, vol. 7, p. 61.

[31]E. Condillac, *Oeuvres philosophiques*, Paris, 1827, pp. 325 - 77.

[32]Poncelet, in *Annales de Gergonne*, 1819.

[33]Gergonne, in *Annales de Gergonne*, 1827.

1.4 Il metodo ipotetico-deduttivo

Sul punto 4 di Enriques non ci soffermiamo. Sul 5 diremo dopo, ma per Enriques il più importante è il 6.

Dalle geometrie non euclidee e dalle geometrie iperspaziali sorge "il concetto generale di una *geometria astratta capace di ricevere diverse intepretazioni*. Virtualmente questo concetto sta già nella trattazione di Plücker e nelle sue coordinate".

Di qui, commenta Enriques, "la comparazione diretta di due ordini di proprietà geometriche, o di due geometrie, unificate nella rappresentazione analitica, conduce più avanti, invitando a *tradurre l'una nell'altra diverse forme di intuizione* ... In questo concetto è contenuta in germe la più larga estensione del principio di dualità, come 'principio delle infinite interpretazioni possibili di una geometria astratta' ".

In cosa consista il "nuovo concetto della scienza dimostrativa"[34] si può dire in modo semplice.

(a) Non si richiede più nessuna definizione reale, ma solo definizioni nominali, quindi alcuni concetti primitivi devono essere non definiti.

(b) I postulati sono enunciati come tali, nella forma di puri rapporti logici supposti tra i concetti primitivi.

La presentazione non sembra gravida di conseguenze, sembra quello che si è sempre normalmente inteso con metodo assiomatico, ma bisogna prestare attenzione ai termini usati, e a cosa si nasconde dietro "la forma di puri rapporti logici". Nell'esame di un sistema per la geometria proposto da Hoüel basato sul concetto di movimento, Enriques commenta:

"Se sono dati dei concetti primitivi $A, B, C, \ldots$, un postulato pone fra di essi una certa relazione $\varphi(A, B, C, \ldots)$: proviamo a tradurla, chiedendoci se sia vera o falsa per altre interpretazioni di $A, B, C, \ldots$ In generale *la traduzione non ha senso se la detta relazione fa appello direttamente al significato intuitivo di* $A, B, C, \ldots$ Come tradurremo, per esempio, i principi di Hoüel prescindendo dal significato specifico del 'movimento' come operazione fisica sulle figure?

La *forma logica* che si vuol dare ai postulati è precisamente quella di *relazioni aventi un significato indipendente dal particolare contenuto dei concetti*, cioè di relazioni affatto generali che possono sussistere fra 'enti astratti' ...

[34]Si veda Enriques, cit., Cap. III, paragrafo 29.

Si può dire che 'infiniti ordini di proprietà geometriche relative ad enti del nostro spazio euclideo, possono ritenersi come interpretazioni di una geometria non-euclidea o anche di una geometria a più di tre dimensioni'. Così nell'esempio di Beltrami, la geometria non-euclidea di una superficie piana si riflette nella geometria delle figure curvilinee tracciate sopra una superficie di curvatura negativa, dove la linea geodetica prende il posto della retta. E, secondo Klein, il sistema di rette dello spazio ordinario ci offre l'immagine di una varietà di second'ordine a quattro dimensioni, immersa in uno spazio lineare a cinque dimensioni[35].

Appunto con Klein e Lie il concetto della geometria astratta ha ricevuto un grande sviluppo, divenendo poi (dopo Segre) un ordinario istrumento di lavoro ... Infatti nulla è più fecondo che la moltiplicazione dei nostri poteri intuitivi recata da codesto principio: pare quasi che agli occhi mortali, con cui ci è dato esaminare una figura sotto un certo rapporto, si aggiungano mille occhi spirituali per contemplarne tante diverse trasfigurazioni ... "

Enriques mette dunque l'accento sulla possibilità, la desiderabilità e quasi la necessità di molteplici interpretazioni come significato o contropartita dei "puri rapporti logici", anche se viene ancora riconosciuto un particolare contenuto dei concetti, che tuttavia non deve comparire negli assiomi.

Il sistema dei postulati porge la definizione implicita dei concetti primitivi, come un sistema di equazioni limita il campo di variabilità delle incognite.

Anche Poincaré accetta le definizioni implicite per la geometria, benché ponga un limite, ad esempio le rifiuta come fondamento dell'aritmetica: "certi assiomi indimostrabili della matematica non sarebbero che definizioni mascherate. Questo punto di vista è spesso legittimo; io stesso l'ho ammesso per quanto riguarda ad esempio il postulato di Euclide. Gli altri assiomi della geometria non bastano per definire completamente la distanza; la distanza sarà allora, per definizione, tra tutte le grandezze che soddisfano gli altri assiomi, quella che è tale da rendere vero il postulato di Euclide"[36].

[35] Le rette dello spazio proiettivo $\mathbb{P}^3$, ovvero di $\mathbb{C}^4$, sono individuate da due vettori $\langle a_0 : a_1 : a_2 : a_3 \rangle$ e $\langle b_0 : b_1 : b_2 : b_3 \rangle$ che danno origine alla matrice

$$\begin{bmatrix} a_0 a_1 a_2 a_3 \\ b_0\, b_1\, b_2\, b_3 \end{bmatrix}$$

i cui sei minori sono le coordinate di Plücker della retta, elemento di $\mathbb{P}^5$. I sei minori non sono indipendenti, e si scende a quattro dimensioni. Le relazioni tra sei numeri perché siano i minori di una simile matrice si esprimono con un'equazione di secondo grado.

[36] H. Poincaré, *Science et méthode*, Flammarion, Paris, p. 161.

Sul fatto che gli assiomi non descrivono una realtà anche Poincaré concorda, sia pure con una sua posizione originale, il convenzionalismo, che ha dato lo spunto a molte discussioni.

"Gli assiomi non sono giudizi analitici *a priori*, sono convenzioni ... La geometria non è una scienza sperimentale; l'esperienza non è che l'occasione per noi per riflettere su idee geometriche che preesistono in noi ... La nostra scelta non è imposta dall'esperienza, ma semplicemente guidata dall'esperienza. Ma resta libera: noi scegliamo questa geometria piuttosto che quest'altra non perché è più *vera*, ma perché è più *comoda*.

[...] Trasportati su un altro mondo, noi potremmo senza dubbio avere una geometria differente, non perché la nostra geometria avrebbe cessato di essere vera, ma perché sarebbe divenuta meno comoda di un'altra"[37].

Il primo lavoro in cui si adotta la presentazione rigorosa di un sistema ipotetico-deduttivo per la geometria è a giudizio di Enriques il trattato di Moritz Pasch[38], seguito da Peano e da altri.

"Per quanto ci è dato di giudicare, ricordando, il senso della forma logica dovette essere riguadagnato, come una *conquista personale*, forse da ognuno dei critici matematici appartenenti alla stessa generazione; sebbene non si possa escludere, in maniera assoluta, un'influenza generica, più o meno diretta, dei predecessori". In particolare "si deve ritenere che ancora abbia avuto a riguadagnarle per proprio conto David Hilbert ..."

Lo hanno riguadagnato nel lavoro sul campo, per il superamento dei loro problemi. Hilbert non voleva affatto costruire un monumento al metodo ipotetico-deduttivo, come la storia ha deciso; voleva studiare l'effetto dei vari assiomi, e di teoremi come quello di Pascal, e rendere indipendente la geometria dai numeri.

La maggior parte dei lavori che hanno costituito la "critica dei principi" non erano un esercizio in assiomatizzazione ma miravano a delineare bene i confini della geometria proiettiva (Pasch, Pieri). Il che testimonia comunque che prima del loro lavoro la geometria proiettiva non era svolta nell'ambito di un sistema assiomatico, questione sulla quale torneremo[39].

Ricordiamo l'esordio delle *Grundlagen del Geometrie* di David Hilbert, 1898. Hilbert non premette alla sua essenziale esposizione alcuna considerazione metodologica, ma esordisce con una definizione:

---

[37]H. Poincaré, "On the Foundations of Geometry", *The Monist* IX (1898), pp. 1-43.

[38]M. Pasch, *Vorlesungen über eine neuere Geometrie*, Lipsia, 1882.

[39]E le contese sui metodi da usare si cristallizzano nella contrapposizione tra metodi sintetici e metodi analitici, o Steiner-von Staudt *vs* Möbius-Plücker.

"Si considerino tre distinti insiemi di oggetti. Gli oggetti del primo insieme siano chiamati *punti* ...; gli oggetti del secondo insieme siano chiamati *rette* ...; gli oggetti del terzo insieme siano chiamati *piani* ...

Tra i punti, le rette e i piani si considera che sussistano certe reciproche relazioni e queste relazioni sono denotate da parole come "giacere", "tra", "congruenti". La descrizione precisa e matematicamente completa di queste relazioni segue dagli assiomi della geometria".

Poincaré, che a proposito dell'opera di Hilbert dirà "Ecco un libro di cui penso molto bene, ma che non raccomanderei a un liceale"[40], così commenta:

"Quello che ci colpisce subito nella nuova matematica è il suo carattere puramente formale: 'Pensiamo, dice Hilbert, tre generi di cose che chiameremo punti, rette e piani, conveniamo che una retta sarà determinata da due punti e che invece di dire che questa retta è determinata da quei due punti noi potremmo dire che ella passa per i due punti o che i due punti sono situati su tale retta'. Che cosa siano queste *cose*, non solamente non lo sappiamo, ma non dobbiamo neanche cercare di saperlo. Non ne abbiamo bisogno, e una persona che non avesse mai visto né punti, né rette, né piani potrebbe fare della geometria bene quanto noi. Che la parola *passare per* o la parola *giacere su* non provochino in noi nessuna immagine, la prima è semplicemente sinonimo di *essere determinata* e la secondo di *determinare* ...

Questo carattere formale della sua geometria, io non lo rimprovero a Hilbert. Era ciò a cui doveva tendere, dato il problema che si era posto. Egli voleva ridurre al minimo il numero degli assiomi fondamentali della geometria e farne l'enumerazione completa; ora, nei ragionamenti nei quali il nostro spirito resta attivo, in quelli nei quali l'intuizione gioca ancora un ruolo, in questi ragionamenti vivi, per così dire, è difficile che non si introduca un assioma o un postulato che passino inavvertiti. Solo dopo aver ricondotto tutti i ragionamenti geometrici a una forma puramente meccanica, egli ha potuto essere certo di essere riuscito nel suo intento"[41]

Con questo commento Poincaré mette in luce l'aspetto formale, e sposta il problema a come si ragiona.

Enriques non si era lasciato distrarre, da questa accusa o rischio di formalismo, da ciò che riteneva importante. Postasi la domanda "se la geometria non-euclidea o la geometria iperspaziale (indipendentemente dalla possibilità

---

[40]H. Poincaré, "Les définitions générales en mathématiques", *L'Enseignement mathématique*, VI, 1904, pp. 257-83.

[41]H. Poincaré, *Science et méthode*, p. 156.

metafisica che taluno può scorgervi) sia un puro schema di formule algebriche" aveva risposto con l'osservazione già sopra riportata che "infiniti ordini di proprietà geometriche relative ad enti del nostro spazio euclideo possono ritenersi come interpretazioni di una geometria non-euclidea o anche di una geometria a più di tre dimensioni"; egli metteva in luce cioè la ricchezza di interpretazioni associata al formale.

Altri ribadiranno questo aspetto, ciascuno con le sue particolari preoccupazioni ed enfasi. Mario Pieri[42] ad esempio è sulla linea di Enriques:

"La caratteristica più importante delle cose primitive di ogni sistema ipotetico-deduttivo è d'essere suscettibili di interpretazioni arbitrarie, entro certi limiti che sono indicati dalle proposizioni primitive ..."

Beppo Levi[43] aggiunge altre sfumature: "Un simbolo rappresenta un'idea primitiva sempre e solo quando è indeterminato il significato che gli compete: rappresenta un'idea derivata quando il suo significato resta individuato tosto che siano fissati i significati delle idee assunte come primitive". E insiste sulla "indeterminazione" del significato delle idee primitive: "è ben vero che un sistema dato di postulati può dare di una idea primitiva una determinazione, in rapporto ad altre idee, minore di quella che effettivamente si attribuisce a quel nome nel discorso comune: ma la vera e completa determinazione di una idea primitiva non è possibile, comunque complesso sia il sistema dei contrassegni che per essa si vogliano enunciare; noi non potremo mai identificare le idee, ma potremo solo affermare che tra esse sussistono certe relazioni".

Per Alessandro Padoa[44] invece la molteplicità di interpretazioni è solo una possibilità: "Può capitare che ci siano diverse interpretazioni del sistema dei simboli non definiti che verificano il sistema delle proposizioni non dimostrate, e dunque tutte le proposizioni di una teoria. Il sistema di simboli non definiti può allora essere considerato come l'*astrazione* ottenuta da tutte queste interpretazioni, e la *teoria generica* ... come l'astrazione di tutte le

[42]M. Pieri, "I principii della geometria di posizione composti in un sistema logico deduttivo", *Memorie R. Accad. Scienze di Torino*, ser. 2a, 48, 1899, pp. 1-62. Tra i teorizzatori del metodo assiomatico citiamo ora soprattutto geometri, e allievi della scuola di Peano, ma se ne potrebbero ricordare anche dal campo algebrico, soprattutto negli USA; alcuni saranno menzionati in seguito.

[43]B. Levi, Antinomie logiche, 1908, in *Opere scelte*, a cura dell'UMI, Cremonese, Roma, 1999.

[44]A. Padoa, "Essai d'une théorie algébriques des nombres entiers, précédé d'une introduction logique à une théorie déductive quelconque", *Bibl. Congrès Intern. de Philos., Paris, 1900*, A. Colin, Paris, vol. 3, 1900, pp. 309-65.

teorie specializzate che risultano dalla teoria generica per la sostituzione successiva dei simboli non definiti con ciascuna delle interpretazioni di questa teoria."

Per parte sua Pasch difende invece proprio il carattere formale delle dimostrazioni, riecheggiando gli algebristi inglesi:

"Occorre in effetti, perché la geometria diventi veramente una scienza deduttiva, che il modo nel quale si tirano le conseguenze sia ovunque indipendente dal *senso* dei concetti geometrici, come deve esserlo dalle figure; sono da prendere in considerazione solo i *rapporti* tra i concetti geometrici, posti dalle proposizioni e dalle definizioni adottate. Durante la deduzione, è certo permesso, ed è utile, pensare al significato dei concetti geometrici considerati, ma non è in alcun modo necessario; tanto è vero che è proprio quando diventa necessario che si manifesta il carattere lacunoso della deduzione e (quando non si può sopprimere le lacune modificando il ragionamento) l'insufficienza delle proposizioni invocate come strumento di prova".

Il requisito di cui parla Pasch discende dallo scopo prefisso: "Se si è dedotto ... un teorema da un gruppo di proposizioni - che chiameremo generatrici (Stammsätze) - il valore della derivazione sorpassa il suo scopo iniziale. Perché se si tira, da proposizioni generatrici, delle proposizioni corrette, cambiando i concetti geometrici con altri [...] si ottiene, senza ripetere la deduzione una proposizione [...] che è conseguenza delle trasformate delle proposizioni generatrici"[45].

Succede ora che le parole della lingua naturale, se usate, tendono a importare con sé nel linguaggio scientifico vecchi ricordi di relazioni usualmente implicate nelle abitudini degli scambi di pensiero in lingua corrente, e bisogna epurarne la scienza: "Anche se nessuna immagine sensibile è ammessa, e neanche una rappresentazione mentale di una tale immagine, l'uso di numerose parole, con le quali sono descritti i concetti geometrici più semplici esercita già in sé una certa influenza"[46].

Ne segue che l'uso della logica formale, o almeno il riferimento alla sua impostazione, diventa essenziale, come aveva intuito Poincaré, e come altri riconosceranno: se è vero che i sistemi deduttivi vanno riferiti a qualche dominio di conoscenze razionali o empiriche, tuttavia la perfezione ideale si ha se le conclusioni relative ai sistemi deduttivi sono stabilite per pura logica; oppure

[45] Pasch, cit.

[46] *Ibidem*.

il dominio può essere aritmetica o geometria, ma "senza ricorrere ad altro sistema ausiliario, di cui possa mettersi in dubbio l'esistenza matematica"[47].

[47]M. Pieri, "Circa il teorema fondamentale di Staudt e i principi della Geometria Proiettiva", *Atti regia Accad. Scienze di Torino*, XXIX, 1904, pp. 313 - 31.

2 Da Euclide a Hilbert

Adesso torniamo a Euclide per capire dal confronto la portata della "riforma della logica, maturata nei secoli", descritta da Enriques. Quindi percorriamo a volo di uccello alcune tappe della storia della matematica nelle quali si vede il disagio dovuto al rispetto, o al mancato rispetto dei vincoli imposti dal modello artistotelico-euclideo. Si tratta solo di una serie di istantanee, senza alcuna pretesa né di completezza né di rigore storico, in quanto prevale l'interpretazione, finalizzata al tema del nostro discorso.

2.1 Gli *Elementi*

Di Euclide, come è noto, si sa poco o nulla. Vissuto intorno al 300 a. C., al tempo del primo Tolomeo (306-286), fondatore di una scuola ad Alessandria. I suoi meriti, secondo Proclo, sono che "mise insieme gli *Elementi*, raccogliendo molti dei teoremi di Eudosso, e anche portando a dimostrazioni incontestabili molte cose che erano state dimostrate in modo solo approssimativo dai predecessori"[1]. Dei tredici libri degli *Elementi*, quello da 1 a 4 e il 6 trattano la geometria piana, i libri 11-12-13 la geometria solida. Il 5 e il 10 sono sulle grandezze e i rapporti. I libri 7-8-9 sono sull'aritmetica.

Gran parte del contenuto era l'eredità della matematica dei due secoli precedenti. L'assiomatizzazione, anche inseguito, si realizza solo dopo che si è accumulato un ricco corpo di conoscenze.

Ricordiamo alcune delle spettacolari acquisizioni di questo periodo pre-Euclideo. Questa parte dell'esposizione è peraltro una divagazione, non necessaria, che si può saltare.

Dopo gli inizi legati ai nomi di Talete (624-547) e Pitagora (572-497) e ai loro viaggi in Egitto, i geometri greci avevano sviluppato le loro ricerche intorno ad alcuni problemi, tra i quali la quadratura del cerchio, la duplicazione del cubo e la trisezione dell'angolo[2], risolvendoli tutti - non con riga e compasso.

I problemi erano classificati a seconda degli strumenti utilizzabili per risolverli: i problemi *piani* richiedevano solo rette e cerchi; i problemi *solidi* richiedevano le

[1]Per i riferimenti alla storia antica, e le citazioni, quando non sia diversamente segnalato il rinvio è a Sir Thomas Heath, *A History of Greek Mathematics*, 1921, ed. Dover, New York, 1981, 2 voll., e alla edizione curata da Heath dei tredici libri degli *Elementi* nella stessa edizione Dover.

[2]Si potrebbe parlare di veri e propri programmi di ricerca, nella terminologia della filosofia della scienza.

coniche, ottenute come sezioni di un cono; i problemi *lineari* richiedevano curve più complicate, come spirali, concoidi, cissoidi.

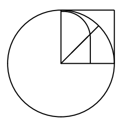
I greci erano consapevoli che nessuno dei tre problemi sopra menzionati era del tipo piano.

Antifone di Atene, un Sofista dei tempi di Socrate, aveva impostato la quadratura del cerchio con un metodo che anticipava quello di esaustione, pensando che sarebbe arrivato a inscrivere un poligono con i lati così piccoli che il poligono sarebbe coinciso con la circonferenza; chi ha lasciato una descrizione (Simplicio, Eudemo) del suo metodo lo ha criticato dicendo che violava il principio che una retta tocca una circonferenza in un solo punto, o quello della divisibilità all'infinito. Secondo Aristotele non era nemmeno il caso di confutarlo, perché i geometri devono confutare solo le fallacie che sono false deduzioni da principi ammessi, mentre quelle che contraddicono tali principi non vanno prese neanche in considerazione. Pare che una conseguenza dei paradossi di Zenone sia stata quella di provocare una ripulsa degli infinitesimi, con conseguenze che si faranno sentire anche ai tempi di Archimede.

Per la trisezione dell'angolo si usavano curve come la quadratrice, la concoide, o perfino nello spazio curve ottenute come intersezione di superfici. Si cercavano e si costruirono anche metodi meccanici per trovare la soluzione, uno attribuito addirittura a Platone, che invece rigettava l'uso di queste soluzioni dalla geometria.

La quadratrice di Ippia, forse contemporaneo di Eudosso (la curva male indicata in figura interna alla circonferenza) si ottiene dal vertice in alto a sinistra del quadrato immaginando un punto che si muova sulla circonferenza e la sua proiezione sul lato superiore del quadrato che contemporaneamente si abbassa.

La trisezione dell'angolo con la quadratrice



era basata sulla relazione

$$\frac{\rho \sin \varphi}{a} = \frac{\varphi}{\pi/2}$$

dove a è il lato del quadrato, φ l'angolo spazzato dal raggio, ρ la lunghezza del segmento del raggio dal centro alla quadratrice.

La quadratura del cerchio si pensava di ottenerla in base alla preliminare rettificazione della circonferenza indotta dalla proporzione: l'arco di cerchio nel quadrato sta al lato del quadrato come questo sta al segmento che dal centro arriva al punto dove la quadratrice incontra il lato. Solo che per determinare questo punto si doveva approssimarlo, con un procedimento che era in pratica l'approssimazione di π.

Ippocrate di Chio (450-430) aveva ridotto il problema della duplicazione del cubo, che tutti fanno risalire a problemi religiosi, con varie storie, a trovare due medi proporzionali continui tra due numeri dati, riduzione ispirata forse per analogia dal fatto per per duplicare un quadrato occorre trovare un medio proporzionale tra a e $2a$.

Se infatti

$$\tfrac{a}{x} = \tfrac{x}{y} = \tfrac{y}{b}$$

allora

$$\begin{array}{rcl} x^2 & = & ay \\ y^2 & = & bx \\ xy & = & ab \end{array}$$

e si ha

$$\tfrac{x^2}{y^2} = \tfrac{ay}{bx}$$

da cui

$$\tfrac{x^3}{y^3} = \tfrac{a}{b}$$

e se $a = 2b$ allora $x^3 = 2y^3$.

Si vede quindi che si devono considerare coniche che diversi autori, tra cui Archita, Eudosso e Menecmo ottenevano come proiezioni dal cono, o con altre sofisticate costruzioni nello spazio.

A Ippocrate è attribuita la prova che i cerchi stanno tra loro come i quadrati dei diametri, proposizione che in Euclide è dimostra con il metodo di esaustione, che forse quindi era stato anticipato da Ippocrate, se si ammette che Ippocrate facesse dimostrazioni rigorose.

C'erano poi molti risultati su aree e volumi, ad esempio che la piramide è un terzo del prisma e il cono un terzo del cilindro, la quadratura delle lunule, gli irrazionali.

Platone conosceva gli irrazionali, oltre a quelli di Teeteto (le varie radici quadrate); sapeva che un segmento può essere la somma sia di due razionali sia di due irrazionali, come quando è divisa nel rapporto estremo e medio.

Di Eudosso Proclo dice che ha accresciuto i teoremi sulla sezione, che avevano avuto origine in Platone. Platone ne parla, ma siccome l'argomento è trattato da Euclide in II, che contiene matematica pitagorica, come diremo, si pensa che fosse ben noto prima.

In II si dice che i teoremi sulle sezioni sono comuni a geometria e aritmetica, salvo quello del rapporto estremo e medio, che in effetti è provato irrazionale da Euclide in XIII. 6 (se la retta è razionale, entrambi i segmenti sono irrazionali). La proposizione II. 11 recita: "Tagliare una retta in modo che il rettangolo contenuto nel tutto e in uno dei segmenti è uguale al quadrato sul restante segmento", proposizione ripresa in VI. 30 parlando di dividere una retta nel rapporto estremo e medio. Questo pare confermare che i Pitagorici sapevano risolvere geometricamente l'equazione $a(a-x) = x^2$, ma che la trattazione pitagorica numerica delle proporzioni si limitava ai rapporti razionali.

A Platone Proclo attribuisce l'invenzione del metodo dell'analisi, che non è ragionevole perché esso era usato sistematicamente dai precedenti geometri. Il metodo consiste nel "riporta[re] la cosa cercata a un principio noto".

Su Eudosso (408-355) abbiamo poche notizie, Proclo informa che avrebbe accresciuto "il numero dei cosiddetti teoremi generali", forse quelli che si riferiscono a grandezze di ogni genere, e quindi forse la teoria delle proporzioni.

Per fissare le date, si ricordi questa successione, con i principali protagonisti (ma ne mancano molti): Talete 624-547, Pitagora 572-497, Ippocrate 450-430, Archita 430-365, Teeteto 415-369, Platone 427-347, Eudosso 408-355, Aristotele 384-322, Euclide ca. 300, Archimede 287-212.

Le prime raccolte di *Elementi* iniziano dopo la metà del V secolo, anche se purtroppo di altre esposizioni di questo tipo sono rimasti al più frammenti, che non permettono di capire come fossero impostati. Ippocrate di Chio pare ne sia stato appunto il primo autore, ma non ci è rimasto nulla, se non citazioni che riportano i suoi risultati sulla quadratura delle lunule.

Gli argomenti dei primi due libri degli *Elementi* di Euclide riportano in sostanza la matematica pitagorica; quello del terzo libro, sui cerchi, era noto a Ippocrate.

Secondo alcuni tutti e quattro i primi libri di Euclide sono una risistemazione talvolta ingegnosa di una parte della geometria pitagorica e post-pitagorica, prescindendo da ogni considerazione che coinvolga rapporti irrazionali; la teoria delle proporzioni di Eudosso viene infatti esposta nel quinto libro.

La parte aritmetica dei libri seguenti, sesto settimo e ottavo, era in gran parte nota ai pitagorici, che consideravano vari tipi di medie, e avevano una teoria delle proporzioni applicabile solo a grandezze commensurabili, o doveva essere contenuta in *Elementi di aritmetica* che pare siano esistiti, almeno dai tempi di Archita (430-365) se non prima, ma non pervenuti. Molti risultati sui numeri piani e solidi dell'ottavo e nono libro (quadrati e cubi) erano anche noti prima. Il libro decimo con la trattazione degli irrazionali era in gran parte coperto dal lavoro di Teeteto (415-369). Parte del libro undicesimo, sui solidi regolari, era noto, come gran parte del tredicesimo, di nuovo a Teeteto. Il dodicesimo tratta del metodo di esaustione e così sistematico non era apparso in precedenti *Elementi*, anche se qualche risultato era noto. Praticamente di veramente nuovo rispetto a quanto era noto a Platone c'era la teoria di Eudosso.

Il merito matematico di Euclide è stato quello di presentare questa teoria e il metodo di esaustione, ma non è questo che gli ha dato la fama. La fama gli è venuta, nei secoli, dall'esposizione dei primi quattro libri, che hanno costituito il modello e l'ideale delle teorie matematiche.

L'organizzazione assiomatica della geometria potrebbe essersi imposta proprio durante la vita di Platone; Platone infatti passa dall'entusiasmo iniziale per il procedimento dei geometri alla critica del loro modo di blindare i risultati in un corpo chiuso; nei primi dialoghi apprezza il metodo dell'analisi, il ricondurre un problema o un enunciato a problemi o enunciati più semplici che si sanno risolvere o decidere; negli ultimi scritti rinfaccia ai matematici proprio la limitazione, rispetto alla dialettica, di fermarsi agli assiomi nella risalita all'indietro della ricerca della conoscenza.

"Essi [coloro che si interessano di geometrie e calcoli] prendono per dati il pari e il dispari, figure, tre specie di angoli, e altre cose imparentate con queste in ciascun soggetto; assumendo queste cose come note, le prendono come ipotesi e di lì in avanti non si sentono più chiamati a dare qualsiasi spiegazione riguardo ad esse né a sè né agli altri, ma le trattano come chiare a tutti; basandosi su queste ipotesi, procedono direttamente a svolgere il resto dell'argomento finché arrivano, con il consenso generale, alla particolare conclusione alla quale la loro ricerca era finalizzata"[3]

La dialettica invece prosegue all'indietro, o all'insù, oltre gli assiomi. Qualcuno dei geometri dei tempo aveva quindi ritenuto di dover codificare alcuni punti di partenza; purtroppo non abbiamo testimonianze di que-

[3] *Repubblica*, vi, 510 C-E. Continua con il famoso passo nel quale osserva che usano figure, ma che non pensano ad esse bensì alle cose che rappresentano, che sono assolute, e possono essere viste solo con il pensiero, non quelle disegnate.

sto epocale momento, se non per i suoi riflessi sulla teoria della scienza di Aristotele.

Siccome Aristotele (384-322) è di poco anteriore a Euclide, dal confronto di vede quello che c'è di nuovo negli *Elementi* rispetto alle conoscenze che Aristotele attingeva dai geometri a lui contemporanei. Aristotele fa probabilmente riferimento all'immediato predecessore di Euclide, come autore di *Elementi*, certo Teudio.

Aristotele riporta, per illustrare passaggi logici, diverse dimostrazioni geometriche, e alcune differiscono da quelle di Euclide. Questo prova che Euclide non si è limitato a raccogliere risultati esistenti, ma ne ha rivisto la dimostrazione per inserirli nel suo percorso.

Un esempio importante[4] è il caso del *pons asinorum*, o teorema di Talete sui triangoli isosceli, proposizione I.5. Aristotele ne presenta una dimostrazione, per illustrare il fatto che in un qualsiasi sillogismo una delle due premesse deve essere affermativa e universale; ma la sua dimostrazione considera il triangolo iscritto in un cerchio, quindi usa proposizioni che riguardano angoli misti, angoli che Euclide non considera che in rarissimi casi, e in particolare l'uguaglianza degli angoli formati da una corda e dalla circonferenza. Se proposizioni del genere erano prima di Euclide preposte alla dimostrazione di quella che in Euclide è la proposizione I.5, secondo alcuni commentatori l'unica dimostrazione possibile era con qualche forma di sovrapposizione.

Peraltro Aristotele cita anche risultati che non sono negli *Elementi*: la somma degli angoli esterni di un poligono uguale a quattro retti, che pure è una proposizione pitagorica; la caratterizzazione del cerchio come luogo dei punti la cui distanza da due punti dati ha rapporto costante $\neq 1$; la circonferenza come curva di lunghezza minima che racchiude un'area data; solo la piramide e il cubo possono riempire lo spazio.

Aristotele era anche al corrente del metodo di esaustione di Eudosso, benché con atteggiamento critico; alcune sue considerazioni sulla impossibilità di somme infinite hanno forse contribuito a rendere dubbio il procedimento, imponendo in seguito ad Archimede di tornare a difenderlo esplicitamente.

Nel rifiutare l'infinito Aristotele si riferisce esplicitamente ai geometri negando che la restrizione sia loro di impedimento: "infatti, per come stanno le cose, essi non hanno bisogno dell'infinito né lo usano, ma richiedono soltanto che la retta finita sia lunga quanto desiderano".

---

[4]Non è l'unico.

Ma naturalmente l'aspetto più interessante di Aristotele, in questo contesto, riguarda la sua teoria delle scienze dimostrative. Euclide non premette alcuna considerazione metodologica; si nota un sensibile accordo con le indicazioni di Aristotele, che diventeranno canoniche, ma Aristotele stesso potrebbe essersi ispirato nella impostazione e nella terminologia ai geometri precedenti o del suo tempo.

Per Aristotele le scienze dimostrative partono da principi che devono essere indimostrabili, per evitare il regresso infinito, ma comunque devono essere sempre necessariamente veri. Di tutti i principi, gli assiomi sono quelli meglio stabiliti. Con in mente l'esempio della geometria, Aristotele chiama *assiomi* o principi le nozioni comuni, che sono verità autoevidenti. Il *genere*, o argomento, di una scienza è qualcosa che deve essere spiegato dalle *definizioni*: la grandezza nel caso della geometria, l'unità nel caso dei numeri. L'esistenza degli enti indagati doveva essere assunta. In particolare si assume in *ipotesi* specifiche l'esistenza di alcuni enti primari, ad esempio rette e punti, mentre l'esistenza di tutto il resto si dimostra.

Le considerazioni di Aristotele sui principi delle scienze non sono sempre coerenti; in una versione molto seguita, gli assiomi sono i principi comuni a tutte le scienze, mentre i postulati sono quelli tipici di una specifica disciplina.

In altri passi, i postulati sono trattati in modo diverso, distinguendoli addirittura dalle ipotesi. Ipotesi e postulati, in quest'altra prospettiva, riguardano fatti che chi insegna non dimostra, benché in sé siano materia di prova. Nel contesto dialogico dell'insegnamento, l'ipotesi è quella che si assume anche senza sapere se è vera, ma con l'assenso del discente (come potremmo pensare in una dimostrazione per assurdo); il postulato è un'assunzione che a differenza dell'ipotesi non ha necessariamente l'assenso dell'interlocutore, e può addirittura essere creduta falsa.

Gli *Elementi* contemplano nozioni comuni, postulati e definizioni, oltre poi a teoremi e problemi. Il verbo per "postulare" è $\tau\acute{\iota}\theta\eta\mu\iota$.

L'*incipit* è il seguente:

Definizioni

1. Un punto è ciò che non ha parti.
2. Una linea è una lunghezza senza larghezza.
3. Le estremità di una linea sono punti.

4. Una linea retta[5] è una linea che giace in modo uguale rispetto ai punti su di essa.

5. Una superficie è ciò che ha lunghezza e larghezza soltanto.

. . .

8. Un angolo piano è l'inclinazione l'una sull'altra di due linee in un piano che si incontrano e non giacciono sulla stessa retta.

9. E quando le linee contenenti l'angolo sono rette, l'angolo è detto rettilineo.

. . .

13. Una frontiera è ciò che è l'estremità di qualcosa.

14. Una figura è ciò che è contenuto da una frontiera o più frontiere.

15. Un cerchio è una figura piana contenuta da una linea tale che tutte le rette che cadono su di essa a partire da un punto di quelli che giacciono dentro la figura sono uguali tra loro.

16. E il punto è chiamato centro del cerchio.

. . . triangoli e quadrilateri, altre figure e ultima

23. Rette parallele sono rette che, essendo nello stesso piano ed essendo prolungate indefinitamente in entrambe le direzioni, non si incontrano da nessuna delle due direzioni.

Postulati

1. Tracciare una linea retta da un punto a un altro punto.
2. Estendere una retta finita in modo continuo in una retta.
3. Descrivere un cerchio con un centro qualunque e una distanza.
4. Tutti gli angoli retti sono uguali tra loro.

[5] Bisognerebbe propriamente tradurre "segmento", visto che una linea ha estremi, ma abbiamo difficoltà a calarci nel mondo matematico finito dei greci.

5. Se una retta incidendo su due rette produce una somma degli angoli interni da una stessa parte minore di due retti, allora le due rette, se estese indefinitamente, si incontrano dalla parte nella quale la somma degli angoli è minore di due angoli retti.

Nozioni comuni

1. Cose uguali alla stessa cosa sono uguali tra loro.
2. Se uguali sono aggiunti a uguali, i risultati sono uguali.
3. Se uguali sono sottratti da uguali, i resti sono uguali.
4. Cose che coincidono l'una con l'altra sono uguali.
5. Il tutto è maggiore della parte.

Nei libri che trattano altri argomenti, sono premesse ugualmente le appropriate definizioni.

Ad esempio l'inizio del libro V contiene la definizione

5. Quattro grandezze si dicono nello stesso rapporto, la prima alla seconda e la terza alla quarta, quando, se si prende un qualunque equimultiplo comune alla prima e alla terza, e un qualunque equimultiplo comune alla seconda e alla quarta, i primi equimultipli entrambi eccedono, o entrambi sono uguali, o entrambi minori dei secondi equimultipli presi rispettivamente nello stesso ordine,

che è la definizione fondamentale per la teoria delle proporzioni di Eudosso[6].

Non ci sono invece in tutti i restanti libri altri postulati, riguardanti argomenti diversi dalla geometria piana, e tale circostanza è importante per quel che segue.

[6]Nella spiegazione semplificata dovuta a De Morgan, esposta da Heath, quattro grandezze A, B, C e D sono proporzionali se, comunque dati m ed n, se mA giace tra nB e $(n+1)B$ allora mC giace tra nD e $(n+1)D$.

La definizione 5 è letteralmente la definizione di uguaglianza di numeri reali data da Weierstrass. Si vede facilmente la corrispondenza con la definizione di Dedekind, nel senso che se x/y e x'/y' sono lo stesso rapporto (dizione impropria) secondo Euclide, allora individuano la stessa sezione: se a/b è un numero razionale e $a/b < x/y$, allora $ay < bx$, ma quindi per la definizione di Euclide $ay' < bx'$ e $a/b < x'/y'$, e così negli altri casi.

I problemi che si pongono con l'*incipit* degli *Elementi* potrebbero riempire una vita a discuterli, a partire da quelli di interpretazione filologica a quelli matematici.

Ad esempio secondo alcuni le nozioni comuni sono una interpolazione successiva, perché non si capisce perché siano per ultime (oltre ad altre ragioni). Ma nei precedenti trattati di *Elementi* dovevano certamente essere elencate alcune nozioni comuni, perché Aristotele le cita, e in particolare spesso, a scopo illustrativo, la terza di Euclide. Questa è una conferma che prima di Euclide qualche tentativo di organizzazione strutturata era già stato fatto.

Delle nozioni comuni, la quarta sembra avere un carattere geometrico, una specie di definizione di uguaglianza geometrica. Forse svolge il ruolo di sostituto del metodo di sovrapposizione.

Per quel che riguarda i postulati, l'opinione più accreditata è che i primi tre non facciano che tradurre le operazioni dei "tenditori di corde" egiziani.

Il quarto e il quinto hanno a che fare con le costruzioni. Il quarto in verità appare diverso dagli altri: si pensa che non lo si sarebbe potuto provare abbastanza presto se non con una sovrapposizione, e che con la scelta di presentarlo come postulato Euclide volesse evitare questo metodo e nello stesso tempo affermare l'invariabilità delle figure.

Sappiamo da Aristotele che fino ad allora la teoria delle parallele non era stata posta su basi scientifiche: in particolare egli lamenta una *petitio principii* nella trattazione, riferendosi probabilmente alla teoria sviluppata sulla base del concetto di direzione. Sembrerebbe allora che Euclide abbia preso la decisione di esplicitare il principio necessario nella forma di un postulato. Questa mossa ricorda esattamente quello che fece Zermelo con l'assioma di scelta, all'inizio del secolo ventesimo. Quelli che non erano d'accordo cercarono alternative, ad esempio Gemino con una teoria delle rette equidistanti.

Quanto alla corrispondenza con il canone aristotelico, vediamo che è abbastanza stretta, il che può significare che l'elaborazione dei geometri precedenti era piuttosto avanzata, se Aristotele si è ispirato ad essa. Le ipotesi (che dovrebbero garantire l'esistenza) in verità non sono del tutto esplicite. L'esistenza delle rette è data dai postulati 2 e 3, ma quella dei punti è assunta tacitamente. Le definizioni invece sembrano svolgere il ruolo che Aristotele prevede. Le definizioni rappresentano naturalmente la più clamorosa deviazione dall'impostazione moderna (le nozioni comuni possono farsi approssimativamente corrispondere alle regole puramente logiche). Si ricordi come

Enriques insistesse che nella costruzione di un sistema ipotetico-dedutivo non comparisse alcuna definizione reale.

Le prime definizioni illustrano aristotelicamente il *genus*: un punto è ciò che non ha parti, una linea ha una lunghezza senza larghezza, una retta è una linea che giace nello stesso modo con suoi punti[7]. Le successive definizioni delle singole figure hanno invece una funzione operativa, e riguardano proprio le cose non primarie che studia la scienza e la cui esistenza secondo Aristotele deve essere dimostrata. Per il cerchio si ha un postulato apposito, perché il cerchio è una figura primaria, per le altre figure le definizioni preparano la strada per le successive costruzioni[8].

Alcune definizioni, come quelle di rombo e romboide, non sono riprese per nulla nel resto del testo; prova forse che Euclide ha fatto *cut and paste* con altri *Elementi*.

2.2 Aritmetica

Non ci sono postulati per i numeri nei libri degli *Elementi* ad essi dedicati, solo definizioni.

Il libro VII inizia con le definizioni

1. Una unità è ciò in virtù del quale ciascuna delle cose che esiste è chiamata uno.

2. Un numero è una moltitudine composta di unità.

3. Un numero è una parte di un numero, il più piccolo del più grande, quando esso misura il più grande;

4. ma è parti quando non lo misura.

[7]Definizione quasi incomprensibile; l'interpretazione più allettante, ma probabilmente forzata, è quella di Proclo, secondo il quale vorrebbe dire praticamente che ha la stessa lunghezza dei punti che stanno tra i due estremi - quindi che è la linea più breve tra due punti. La lettura prevalente è che la retta si presenta con la stessa forma rispetto a tutti i suoi punti.

[8]L'origine delle prime definizioni è un argomento controverso; quella di angolo come inclinazione sembra comparire qui per la prima volta, e solleva problemi sulla natura dell'angolo. La definizione di retta qualcuno la vede come una riformulazione di quella di Platone, per il quale la retta è "quella di cui il mezzo copre gli estremi"; forse significa che guardandola da un estremo si vede solo l'estremo. Euclide potrebbe averla riformulata in modo da non fare riferimento alla vista e quindi ai sensi.

Prosegue con le definizioni di multiplo, pari, dispari, primo, composto e così via fino a perfetto.

La definizione di numero riprende l'idea di Aristotele che l'aritmetica è la scienza delle unità (così come la geometria è la scienza delle grandezze). Questa e simili definizioni erano comuni nella matematica pitagorica.

La definizione è tuttavia obsoleta in considerazione della scoperta dell'incommensurabilità: se i numeri fossero composti di unità, la diagonale del quadrato dovrebbe essere un multiplo di un'unità di misura scelta come un sottomultiplo del lato. Un'altra tradizione, sovrapposta, considera i numeri come legati alle figure geometriche elementari, anche se inizialmente le figure erano individuate dai sassolini; si distinguono così ad esempio numeri lineari, triangolari, quadrati.

Si noti che per quel che riguarda la teoria delle proporzioni per grandezze, nel libro V, si può pensare che le definizioni in parte svolgano il ruolo di postulati, in quanto incorporano alcune assunzioni. Quella delle grandezze commensurabili tra loro (definizione 4) impone comunque una restrizione a un certo tipo di grandezze, e si può pensare come un sostituto del lemma di Archimede (si veda oltre):

4 Due grandezze si dicono avere un rapporto, l'una all'altra, se ciascuna di esse, se moltiplicata, può eccedere l'altra.

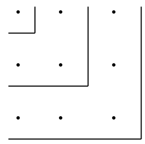
Euclide riservava di fatto il nome di grandezze alle grandezze che ora si chiamano archimedee, pur conoscendone altre. In III, 16 infatti Euclide mostra che non esiste alcun sottomultiplo di un angolo rettillineo che sia minore dell'angolo di contingenza (angolo curvilineo tra circonferenza e tangente). La proprietà archimedea è usata ad esempio in V. 8., che afferma che di due grandezze diverse, la più grande ha con una terza qualsiasi un rapporto maggiore di quella che con essa ha la più piccola.

Nelle definizioni per i numeri interi positivi non c'è neanche nulla di simile, se non il fatto che tra i numeri può sussistere la relazione di divisibilità.

L'assenza di principi espliciti non aveva impedito uno sviluppo sorprendente dell'aritmetica: in Euclide si trovano l'algoritmo per il massimo comune divisore e il teorema sull'infinità dei numeri primi (IX. 20); erano note le formule per la somma dei primi n numeri e quella per la somma dei quadrati dei primi n numeri.

La tecnica principale era rappresentata dai procedimenti iterativi basati sullo gnomone.

Nella seguente figura si vedono tutte le caratteristiche dell'aritmetica prearchimedea: la rappresentazione dei numeri come distribuzione di punti nel piano, la loro classificazione secondo la forma della figura (in questo caso ad esempio il quadrato), l'uso dello gnomone per realizzare processi ricorrenti. Lo gnomone, in origine una squadra, evolve a essere definito da Erone di Alessandria come quello che, aggiunto a una figura o a un numero li rende simili a quello a cui è stato aggiunto.



Dalla figura si capisce come si arrivi alla formula per la somma dei primi n numeri dispari, e come il risultato sia un quadrato[9].

Archimede dimostrerà tali risultati senza usare l'ausilio geometrico dello gnomone, e andrà oltre, usando altri artifici.

Ad esempio, per ottenere la formula per $1 + 2 + \ldots + n$ Archimede scrive

$$\begin{array}{c} 1 + 2 + \ldots + (n-1) + n \\ n + (n-1) + \ldots + 2 + 1 \end{array}$$

e nota che sommando per colonne si ha $n + 1$ per n volte, da cui il risultato

$$\tfrac{n(n+1)}{2}.$$

L'artificio è stato attribuito a Gauss, nella variante di sommare direttamente primo e ultimo addendo, secondo e penultimo, e così via.

Se si esaminano i calcoli di Archimede si vede quali proprietà dei numeri sono sfruttate; nella nostra terminologia: la commutatività della somma, la proprietà associativa della stessa, il fatto che la somma di k addendi uguali ad a è il prodotto ka (cioè che la moltiplicazione è la somma iterata), che un'uguaglianza resta tale se si moltiplicano ambo i membri per lo stesso numero, e che ogni numero ha un inverso (o almeno lo possiede il numero 2,

[9]Per altre applicazioni dello gnomone, si veda P. Zellini, *Gnomon*, Adelphi, Milano, 1999.

per scrivere l'espressione di sopra, ma allora certamente ogni intero diverso da zero[10]).

Nel libro VII di Euclide non si trovano le dimostrazioni di queste proprietà; solo la commutatività della moltiplicazione è dimostrata e, in altra terminologia, la proprietà che un rapporto resta uguale se si moltiplicano ambo i membri per lo stesso numero. Nel libro VII la trattazione è basata sulle proporzioni, e i multipli e sottomultipli; le proposizioni riguardanti $1/n$ si trovano in questa forma indiretta. Si studia il massimo comune divisore, e il minimo comune multiplo, le proprietà dei numeri primi e dei fattori primi (i numeri sono sempre rappresentati da segmenti). Alcune leggi, come la distributività, erano state dimostrate prima geometricamente (si veda oltre).

L'aritmetica prende avvio dalle proprietà menzionate, e analoghe, non dalla definizione del numero; nello stesso modo inizia ancor oggi l'insegnamento della stessa ai bambini, dopo aver stabilito o accettato nei modi più disparati (fisici, intuitivi) proprietà algebriche di base.

Possiamo chiederci quali altri principi dimostrativi fossero impliciti nei ragionamenti dei greci, oltre alle proprietà già indicate.

Un passo fondamentale per lo sviluppo dell'aritmetica è quello di associare alle operazioni dirette, come la moltiplicazione, quelle inverse; la divisione, che fornisce quoziente e resto, è espressa da quello che oggi si chiama teorema fondamentale della divisione, e afferma che per ogni m ed n, con $m > n$, esistono unici q ed r tali che

$$m = nq + r, \qquad \text{con} \qquad r < n,$$

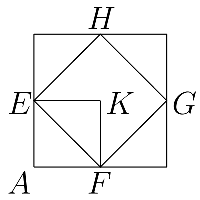
risultato in verità valido anche per $m \leq n$.

Nell'algoritmo del massimo comune divisore (in Euclide VII) si applica ripetutamente la divisione a dividendi sempre minori, finché quindi il procedimento termina quando si arriva a 0. La stessa esecuzione (e dimostrazione della possibilità) della divisione avviene per sottrazioni successive che diminuiscono il dividendo.

Si dà quindi per scontato che una successione decrescente di numeri non prosegue indefinitamente, ma arriva a 0, o a 1. Era attribuita al lavoro dei pitagorici la scoperta che "mentre le grandezze sono divisibili *ad infinitum*, i numeri, divisi, lasciano sempre una parte minima non suscettibile di ulteriore divisione", che non è altro che il moderno principio della discesa finita, equivalente all'induzione.

[10]Oppure in modo meno impegnativo la proprietà che il prodotto di due numeri successivi è pari, quindi divisibile per 2.

Tale principio è alla base della dimostrazione dell'incommensurabilità di lato e diagonale del quadrato, dove a partire da un quadrato con diagonale d e lato l interi si costruisce, supponendo per assurdo $2l^2 = d^2$, un quadrato più piccolo con diagonale l e lato $d/2$ interi.



(Nel quadrato $EFGH$ sia il lato $EF = l$ e la diagonale $FH = d$. Il quadrato $AFKE$ ha lato $d/2$ e diagonale l.)

Questa proprietà era usata in generale senza essere menzionata, e raramente; non sono noti ragionamenti che assomiglino alla nostra forma dell'induzione, in avanti, se non forse un discusso passo di Platone; d'altra parte gli usi del principio di induzione, o meglio forme equivalenti come la discesa finita, compariranno sporadicamente nella storia della matematica, in Fermat ed Eulero ad esempio; Pascal sarà il primo a formulare il principio di induzione nel modo preciso moderno.

Più consapevole, e discusso, è stato l'uso di un altro principio. Nel teorema della divisione, una possibile dimostrazione consiste nel considerare, supposto $n > 1$, i multipli

$$n, 2n, 3n, \ldots$$

finché non si arriva a un k tale che o $kn = m$ o $kn > m$, e in questo caso $k - 1$ è q.

Se si ragiona in questo modo, si assume quindi che dati m ed n, con $1 < n < m$ esiste un primo numero k tale che $kn \geq m$.

Questa proposizione non è altro che il cosiddetto assioma di Archimede[11].

2.3 Archimede

L'opera di Archimede, del terzo secolo a.C., subito dopo Euclide, è particolarmente significativa dal punto di vista fondazionale. In Archimede è chiara

[11] L'assioma considerato nel contesto dei numeri reali asserisce che dati due numeri reali a e b, $a < b$, esiste un n tale che $an \geq b$.

la volontà di assiomatizzare gli argomenti matematici sul modello della geometria, quando vi riesce, come è altrettanto chiara la difficoltà a farlo con argomenti numerici.

Le sue dimostrazioni geometriche delle misure di aree e volumi utilizzavano il metodo di esaustione di Eudosso; il metodo di esaustione era un metodo di approssimazione che richiedeva come coronamento una dimostrazione di convergenza; questa a sua volta consisteva in genere in una dimostrazione per assurdo, assumendo uno scarto tra il risultato del processo e l'obiettivo, e mostrando quindi che qualche approssimazione si avvicinava meno dello scarto. Il metodo era perfettamente corretto, come in seguito si è riconosciuto, in un contesto in cui fossero accettati gli assiomi adatti, in particolare l'assioma che ora porta il nome di Archimede (e che avrà un ruolo importante nella teoria moderna dei numeri reali e della geometria), ma che non è stato inventato da lui.

Archimede ne parla come di un lemma, che prima del significato usuale che ha acquisito significava semplicemente "assunzione"; egli ci dice che i teoremi sulla piramide e il cono, che erano stati scoperti da Democrito, sono dimostrati da Eudosso, usando un lemma che è stato usato anche da precedenti geometri anche per provare che i cerchi stanno tra loro come i quadrati del diametro e le sfere come i cubi; e siccome quello sui cerchi si dice che sia stato dimostrato da Ippocrate, ed è dimostrato da Euclide con il metodo di esaustione, può darsi che il metodo sia stato anticipato da Ippocrate, ma che Eudosso abbia formulato esplicitamente il lemma. Il lemma non si trova in Euclide, ma è in pratica equivalente come si è detto alla definizione 4 del libro V sulle grandezze.

Archimede nel presentare e commentare i propri risultati con i colleghi alessandrini perora la causa dell'accettabilità del principio, ritenendolo sufficientemente dotato di affidabilità ($\pi\acute{\iota}\sigma\tau\iota\sigma$) in quanto usato per risultati già conosciuti e che avevano meritato la *pistis* della comunità.

Archimede dichiara: "A me basta aver condotto i teoremi che ora rendo pubblici con la stessa affidabilità" di quelli di Eudosso.

Le dimostrazioni originarie di questi risultati erano state in verità ottenute da Archimede con il suo metodo meccanico, consistente in pratica nel considerare rapporti di aree proporzionali al rapporto tra il loro peso (come si fa ancora adesso didatticamente con i bambini) ovvero a prendere in considerazione il loro baricentro. In questi calcoli usava l'artificio degli infinitesimi, o dei posteriori *indivisibili* di Cavalieri, considerando una figura come affettata da segmenti.

Proprio ad Archimede, insieme con altri, sembrano dovute le prime formulazioni dei principi della statica; negli *Elementi di meccanica* archimedei si trova di fatto un abbozzo di teoria matematica in presentazione assiomatica. Per i baricentri, alcuni degli assiomi sono i seguenti:

1. Se da una grandezza si toglie un'altra grandezza e se uno stesso punto è il centro di gravità della grandezza intera e di quella tolta, questo stesso punto è il centro di gravità della grandezza restante.

2. Se da una grandezza di toglie un'altra grandezza, e se la grandezza intera e quella tolta non hanno lo stesso centro di gravità, il centro di gravità della grandezza restante si trova prolungando, oltre il centro della grandezza intera, la congiungente i centri di gravità della grandezza intera e di quella tolta e togliendo da questa un segmento che abbia con la congiungente i sopraddetti centri di gravità lo stesso rapporto che ha il peso della grandezza tolta con il peso della grandezza restante.

3. ...

4. Il centro di gravità di una retta è il suo punto di mezzo.

5. Il centro di gravità di un triangolo è il punto in cui si tagliano scambievolmente le rette condotte dai vertici del triangolo al punto di mezzo dei lati.

6. Il centro di gravità di un rettangolo è il punto in cui si incontrano le diagonali,

e altri analoghi; è facile immaginare come si calcola a partire da essi il baricentro di figure poligonali regolari e non regolari.

Queste sono le proprietà usate di fatto da Archimede nel metodo meccanico; ma egli non riesce a integrarle nel quadro dei principi universalmente ammessi e codificati; forse addirittura non riesce a legittimare la nuova scienza con pari dignità accanto a quella delle grandezze e delle unità; è consapevole del fatto che le sue giustificazioni non sono ritenute conclusive, anche se non per difetti intrinseci, o forse a causa degli infinitesimi, e che il metodo è considerato solo una tecnica euristica.

Lui stesso si sente perciò costretto a inventare dimostrazioni diverse, quelle con il metodo di esaustione, che siano dimostrazioni geometriche basate solo su principi accettati.

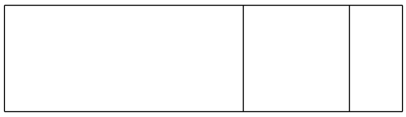
L'attività di Archimede testimonia anche come la formazione dei principi dipenda dallo svolgimento di dimostrazioni, se e quando esse ricevono l'assenso da parte della comunità; la *pistis* era una caratteristica che Platone collegava solo alla $\delta\acute{o}\xi\alpha$, mentre Aristotele la riteneva applicabile sia alla $\delta\acute{o}\xi\alpha$ che all'$\epsilon\pi\iota\sigma\tau\eta\mu\eta$. I matematici la ritenevano essenziale per le dimostrazioni, e discutevano piuttosto tra di loro che non con i filosofi.

Dopo Archimede si ha una stagnazione, e un riflusso, dovuto anche al fatto che pare si siano abbandonati del tutto i metodi geometrici, e soprattutto la rappresentazione geometrica dei numeri; nel trattato di Nicomaco di Gerasa, uno degli ultimi pitagorici, intorno al 100 d.C., quando si devono trattare più numeri insieme si devono usare complicate circonlocuzioni, e in particolare non ci sono più dimostrazioni. Vero è che il testo di Nicomaco sembra che sia un'opera divulgativa, a giudicare dal tono retorico, enfatico, colorato, interessato soprattutto ai risultati sorprendenti e quasi miracolosi. Ma la comparsa di opere del genere è di per sé significativa, allora come oggi.

2.4 Algebra

L'algebra antica e moderna non è per nulla sistematica; a dire il vero non esiste neanche come disciplina autonoma, la sua nascita è molto tarda, lenta e laboriosa. Non che non si conoscano e non si risolvano problemi che noi chiameremmo algebrici, già negli *Elementi* di Euclide, e addirittura presso i Babilonesi, come la ricerca delle terne pitagoriche. Sono in generale equazioni di secondo grado e problemi di analisi indeterminata (meno equazioni di incognite).

Questi problemi nascevano presso i Greci come problemi geometrici ed erano risolti con metodi geometrici - le costruzioni con riga e compasso permettono di risolvere problemi quadratici. Si parla di *algebra geometrica*. Ad esempio, la proprietà distributiva della moltiplicazione rispetto alla somma è espressa, nella proposizione 1 del libro II degli *Elementi*, nella forma che se sono date due rette, e una è divisa in un numero qualunque di segmenti arbitrari, il rettangolo contenuto tra le due rette è uguale ai rettangoli contenuti nella retta non divisa e in ciascuno dei segmenti:



Con le curve come le coniche si risolvevano sistemi di equazioni di secondo grado. Archimede arriva a risolvere anche equazioni di terzo grado; quando poi fiorisce Alessandria ellenistica i numeri e i problemi numerici prendono il sopravvento sulla geometria. Ma il problema era quello della notazione (che nel contesto della presente discussione interessa solo relativamente) sia per le potenze che per incognite e parametri. Con la concezione geometrica del numero si consideravano solo quadrati e cubi.

Diofanto ha inventato una notazione per le potenze fino alla sesta (tre simboli per le prime tre potenze, quindi il loro accostamento per ottenere le altre con la somma degli esponenti); egli usa anche un segno per l'incognita, che è una specie di h rovesciato nei manoscritti, reso in stampa da ζ o una specie di S, probabilmente già utilizzato da altri, che tuttavia purtroppo in seguito si perde.

Il segno è proprio solo una abbreviazione, sicché Diofanto è classificato tra i primi rappresentanti di quella che alcuni chiamano algebra sincopata, rispetto alla precedente algebra retorica, fatta solo a parole e a quella simbolica moderna. L'incognita è definita come da Diofanto "una quantità contenente una moltitudine indeterminata o indefinita di unità" ed è chiamata $\alpha\rho\iota\theta\mu\acute{o}\sigma$, o numero (e qualche volta il segno è usato al posto di "numero" anche quando non è incognito, quindi è davvero una abbreviazione). Il segno dovrebbe essere una contrazione di $\alpha\rho$.

Il testo *Arithmetica* di Diofanto (secondo o terzo secolo d.C., più probabile intorno al 250) non è comunque un trattato ma una collezione di problemi, salvo per una interessante prefazione, di cui diremo in seguito. Si trattano varie equazioni e sistemi di equazioni determinate e indeterminate di secondo grado.

Il lavoro di Diofanto è recuperato dagli arabi, che riprendono i suoi metodi dell'aggiunta dello stesso termine ad ambo i membri (completamento, *al-jabr*, da cui il nome dell'algebra) e della riduzione dello stesso termine da ambo i membri; l'individuazione di mosse del genere testimonia la volontà di dare sistematicità ai metodi di risoluzione.

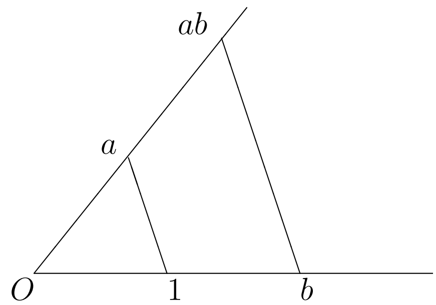
Alla ripresa della matematica in Europa alla fine del Medio Evo, favorita dall'importazione della notazione indiana e delle tecniche arabe, sotto la spinta delle necessità del commercio e della finanza, si sviluppano da una parte il calcolo (Fibonacci, *Liber Abbaci*, 1202) e dall'altra l'algebra.

Dopo Fibonacci bisogna tuttavia aspettare tre secoli prima di trovare Luca Pacioli e poi Bombelli e altri. La notazione è ancora un grave impedimento; si trovano usate o quella di Diofanto, o altre con espressioni artificiose del tipo: "quadrato-quadrato" per la quarta potenza, "primo inesprimibile" per la quinta, "quadrato-cubo" per la sesta.

Non c'è neanche il segno per l'uguaglianza. Per quel che riguarda le espressioni letterali, Luca Pacioli usava ancora n^0 per una costante, *co.* da "cosa" per l'incognita, *ce.* da "censo" per il quadrato dell'incognita, e così via.

Per farla breve, il calcolo letterale simile a quello che usiamo noi deve attendere Descartes nel Seicento. Descartes giudica l'algebra del suo tempo in uno stato di grande confusione.

Descartes riesce a considerare potenze qualunque di un segmento, e quindi andare oltre la terza dimensione a cui erano limitati i Greci con i numeri quadrati e cubi, grazie alla definizione geometrica della moltiplicazione, per mezzo della proporzionalità tra i segmenti tagliati da parallele su due rette incidenti,



proprietà che, per inciso, era stata provata numericamente da Euclide in VII. Tuttavia questa possibilità di generalizzazione non venne esplorata, fino a tardi, come vedremo.

L'algebra antica e moderna non è quindi per nulla sistematica, nonostante i notevoli arricchimenti; i libri, che verranno, hanno ovviamente una loro naturale organizzazione interna, ma di tipo estrinseco, non certo assiomatico; se si studiano le equazioni, si incomincia da quelle di primo grado e poi di secondo e così via - quando non si elencano casi singoli, oggetto di sfide, in mancanza di formule risolutive.

Nel trattato di Eulero[12] ad esempio sono presentati prima metodi di calcolo, cioè le operazioni, per numeri interi, frazioni, irrazionali che risultano da radici, e logaritmi; quindi operazioni su quantità composte, che sarebbero le espressioni algebriche; seguono le progressioni aritmetiche e geometriche, le serie e infine le equazioni dal primo al quarto grado, i sistemi di equazioni e l'analisi indeterminata.

La geometria si arricchisce con la geometria analitica, che presenta evidenti elementi di unificazione nello studio delle curve e superfici, ma anch'essi non di tipo assiomatico.

Un merito della geometria analitica è quello di aver introdotto un granello di polvere che contribuirà a rovinare il meccanismo consolidato: essa mette in crisi l'idea che le scienze siano caratterizzate dal genere che studiano e disobbedisce platealmente al dettato aristotelico che non si possono usare i metodi di un genere nello studio di un altro.

2.5 I numeri negativi

Abbiamo già ricordato come ancora nell'Ottocento i numeri negativi non avessero trovato una sistemazione soddisfacente, e come la loro natura, se ci si pensava, disturbasse i matematici. La difficoltà era proprio fondazionale, e per nulla matematica, a differenza che quella relativa ai numeri immaginari: per questi, che un numero elevato al quadrato dia un risultato negativo contraddice la fondamentale regola dei segni.

Il meglio che riuscivano a fare era il riferimento a metafore, che tuttavia talvolta entrano in contraddizione tra loro. La metafora del "togliere", per la sottrazione, ad esempio non si accorda con quella del nulla misurato dallo zero. "Nello stesso modo in cui i numeri positivi sono incontestabilmente maggiori del nulla, i numeri negativi sono meno di nulla" (Eulero). Parlare di "meno di nulla" è contraddittorio, come chiedersi cosa ci fosse prima del tempo.

[12] Euler, *Anleitung ...*, cit.

Le metafore non coprono tutte le situazioni nelle quali si usa la parola "togliere". Chi non ha nulla può ancora fare un debito, ma se in un sacco non ci sono più mele, non si può più togliere nulla[13].

Pure secondo Eulero come la serie positiva $0, +1, +2, \ldots$ si ottiene "aggiungendo 1 a 0, vale a dire, 1 al nulla; e continuando sempre a crescere così di una unità ... [così] se, invece di continuare questa serie con successive addizioni, la continuassimo nella direzione opposta, sempre sottraendo un'unità, avremmo la serie dei numeri negativi : $0, -1, -2, -3, \ldots$".

Eulero non dà una definizione dei numeri negativi, ma avverte che "Sarebbe certo della massima importanza in tutto lo sviluppo dell'Algebra che ci si formasse un'idea precisa delle quantità negative di cui abbiamo parlato. Io mi accontenterò tuttavia di osservare qui, che tutte le espressioni come

$$+1 - 1, +2 - 2, +3 - 3, +4 - 4, \ldots$$

sono uguali a 0, o al nulla. E che:

$$+2 - 5 = -3:$$

perché se una persona ha 2 corone e ne deve 5 egli non solo non ha nulla, ma deve anche 3 corone".

La metafora dei crediti e debiti, usata anche per giustificare la moltiplicazione, non è l'"idea precisa" auspicata, visto che Eulero si *accontenta* di una presentazione formale.

Se tuttavia ci si chiede che cosa manca rispetto alla definizione che sarà in seguito data, si vede che manca solo di dire che

$$-3 = +2 - 5 = +3 - 6 = \ldots = +12 - 15 = \ldots$$

A suo tempo si useranno gli insiemi al posto dei puntini e una relazione esplicita senza il segno "$-$ per individuare tutte le coppie che vi appartengono:

$$-3 = \{\langle n, m\rangle : n + 3 = m\}$$

raccogliendo in una espressione tutte le uguaglianze

$$2 + 3 = 5, \;\; 3 + 3 = 6, \;\; 3 + 12 = 15, \ldots$$

[13]Ancorché l'insieme vuoto sia ancora un insieme di mele.

implicite sopra.

Si noti che l'uso del linguaggio insiemistico è inessenziale rispetto alla sostanza, che è costituita dalla relazione algebrica $n + 3 = m$, che è quella che si usa nei calcoli e che Eulero mette in evidenza.

Eulero dunque semplicemente si astiene dal dare una definizione (dei numeri negativi), e fornisce invece i criteri per passare dall'operazione di sottrazione tra naturali ai numeri negativi e viceversa.

Tali criteri, se la loro esplicitazione fosse perseguita in modo adeguato, si sostanzierebbero in un insieme di relazioni algebriche tra simboli numerici con prefisso + (soddisfacenti alle leggi dei numeri naturali) e simboli numerici con prefisso $-$, del tipo

$$\begin{array}{lcl} (+m) + (-n) = (-p) & \text{se e solo se} & (+m) + (+p) = (+n) \\ (-m) + (-n) = (-p) & \text{se e solo se} & (+m) + (+n) = (+p) \end{array}$$

e altre analoghe, che nel loro insieme costituirebbero una definizione assiomatica.

Il senso dell'imprecisione o della mancanza, che avevano i matematici, quando non riuscivano a esprimersi in modo soddisfacente nella formulazione dei principi, era dovuto all'idea di dover fare qualcosa che invece non era necessario fare: pensavano che gli assiomi dovessero discendere dalle definizioni, e che quelli che essi usavano come punto di partenza non potessero essere considerati tali se non se ne faceva vedere la discendenza dalle giuste definizioni.

Eulero cercava, per dare una "idea precisa", una definizione che esprimesse il *genus* aristotelico, ma che fosse matematica, a differenza delle metafore che individuavano le modalità dei diversi tipi di numeri con criteri non matematici, ma sociali o fisici.

Proprio a proposito dei numeri negativi, un'indicazione precorritrice si sarebbe potuta trovare in Diofanto. Egli nell'*Arithmetica* chiama "numero" ogni soluzione *positiva* razionale di un problema; le soluzioni dei problemi sono sempre numeri positivi, interi o razionali, mentre quelli negativi intervengono nei calcoli intermedi, esattamente come succederà in seguito con gli immaginari. Se invece un'equazione avrebbe solo soluzioni negative, non viene considerata. Allora Diofanto lavora con i numeri negativi, senza chiamarli tali, nel seguente modo. Considera un'entità che chiama "deficienza" o "mancanza", e formula regole per estendere ad essa le operazioni, esattamente le regole dei segni: "deficienza moltiplicata per deficienza dà disponibilità [numero positivo], deficienza moltiplicata per disponibilità dà deficienza, e il simbolo per la deficienza è la ψ invertita e accorciata" (Def. IX).

Forse anche Diofanto con "deficienza" e "disponibilità" pensava ai debiti e ai crediti, ma non lo dice. Deficienza e disponibilità potrebbero essere considerati alla lettera termini matematici introdotti assiomaticamente. La soluzione di Diofanto si è persa, o comunque non è stata riconosciuta come una soluzione assiomatica. Il processo storico, per ribadire una banalità, non è lineare e cumulativo.

2.6 Gli infinitesimi

La storia del calcolo infinitesimale è abbastanza nota. A differenza che nei precedenti casi considerati, non si ha una matematica comunque rigorosa, anche in assenza di una organizzazione assiomatica, ma una matematica dubbia, e metodi che funzionano come trucchi di magia: gli infinitesimi ora sono diversi da zero ora si pongono uguali a zero; si introducono e si trascurano; le quantità variabili diventano evanescenti, i fantasmi di entità dipartite del vescovo Berkeley; le flussioni sono proporzionali ai momenti, che sono i principi nascenti di quantità finite.

Berkeley protesta che la nostra immaginazione, "che è una facoltà derivata dal senso, è incapace di formarsi una idea chiara delle minime particelle di tempo ... incapace di comprendere i momenti delle quantità fluenti *in statu nascenti*, nella loro prima origine, o esistenza iniziale, prima che diventino particelle finite". Ma "i matematici moderni non considerano le precedenti asserzioni come misteri, bensì come asserzioni chiaramente concepite e dominate dalle loro menti comprensive". Forse è la prima comparsa dell'atteggiamento del matematico che dice di "vedere" i suoi oggetti, con nessuna considerazione delle persone comuni.

Leibniz non si è lasciato invischiare in questa metafisica, e non a caso hanno prevalso i suoi metodi algebrici. Egli pensava che gli infinitesimi fossero elementi ideali, *finzioni* per le quali si doveva supporre che valessero esattamente le leggi dei numeri reali; la loro utilità era innegabile, mentre d'altra parte il metodo di esaustione era sempre disponibile per ridimostrare i risultati senza questi ausili utili ma politicamente scomodi. La tesi di Leibniz era che ogni risultato del calcolo infinitesimale si sarebbe potuto ottenere con il metodo di esaustione.

In termini moderni, si potrebbe dire che nella visione di Leibniz gli infinitesimi fornivano una nuova tecnica dimostrativa, codificata da una teoria (se e quando ci fosse stata una teoria) che era un'estensione *conservativa* della teoria dei numeri reali. Questo significava che nella teoria ampliata si potevano dimostrare - più comodamente - tutti e soli i teoremi non riguardanti gli infinitesimi che si potevano dimostrare nella teoria classica dei reali. Leibniz non si è mai posto tuttavia il problema di precisare quali fossero le leggi che valevano ugualmente per numeri finiti e infinitesimi (anche perché mancava una codifica e una classificazione delle leggi dei numeri reali), né ha osservato che la legge di Archimede non poteva essere una di quelle[14].

Il primo manuale di calcolo differenziale, l'*Analyse des infiniment petits pour l'intelligence des lignes courbes* del marchese dell'Hospital (1696) esibisce la patetica ambizione di ricalcare Euclide. Esso inizia infatti con due definizioni:

Definizione I. Si chiamano quantità *variabili* quelle che aumentano o diminuiscono con continuità; e al contrario [si chiamano] quantità *costanti* quelle che rimangono le stesse mentre le altre cambiano ...

Definizione II. La porzione infinitamente piccola di cui una quantità variabile aumenta o diminuisce con continuità è chiamata la differenza [differenziale].

Si fanno poi delle *demandes* al lettore, nello stile dei postulati che devono essere concessi dall'interlocutore: che sia concesso supporre che si possano identificare due quantità che differiscono di una quantità infinitamente piccola, e che una linea curva possa essere considerata l'*assemblage* di linee rette. Non assiomi, ma al massimo veri e propri postulati in senso antico, che il discente deve accettare pur vedendoli falsi (contraddittori).

2.7 Geometria e realtà

Si è già accennato alle geometrie non euclidee. A conferma di come ancora nel Settecento si pensasse al metodo assiomatico, Saccheri riteneva che "quell'Assioma famoso fu assunto da Euclide come cognito di per sé".

Saccheri è talmente convinto della verità dei postulati della geometria, che nel suo tentativo di dimostrazione la parte riguardante l'ipotesi dell'angolo

[14]Questo lavoro sarà compiuto da Abraham Robinson con l'Analisi non standard, nel 1963.

acuto è conclusa, visto che logicamente non riusciva, con l'affermazione che la proposizione è falsa, in quanto "ripugna alla natura della retta".

Diventa difficile accettare come matematica una teoria con assiomi ripugnanti.

La crisi della geometria in conseguenza della comparsa dei sistemi non euclidei è comprensibile. Con la terminologia di Aristotele, se si modifica un solo assioma, e non le definizioni che fissano il *genus*, che non cambiano, le nuove teorie parlano dello stesso argomento; la realtà descritta è la stessa, sono le descrizioni che sono diverse; come si fa a sapere quale è quella giusta?

E comunque, se una teoria era vera, quale era lo statuto delle altre?

Se non è la natura della retta, esposta nella definizione di Euclide, che è in grado di discriminare le diverse descrizioni, potrebbe sopperire la misurazione dello spazio fisico. Le misurazioni come è noto non riuscirono ad avere effetto dirimente.

Un risultato l'ebbero, quello di far pensare ad alcuni, tra cui Gauss, che la geometria fosse una scienza che tratta di una realtà fuori dalla mente, con inevitabili compromissioni empiriche, e che quindi non fosse parte della matematica pura.

Lo statuto della geometria è svalutato, agli occhi degli stessi geometri; Plücker è uno di quelli che ammette che la geometria viene al secondo posto[15].

2.8 L'aritmetizzazione dell'Analisi

Con questo nome si indica l'impresa collettiva della rigorizzazione dell'Analisi nel corso dell'Ottocento. La necessità deriva dal disagio con i nuovi enti, funzioni. Nessuno sa cosa sia una dimostrazione, diceva Dirichlet. Due erano gli obiettivi principali: il primo la definizione stessa non geometrica del continuo dei reali, ora soprattutto che l'intuizione appariva screditata come fonte di conoscenza di quella che si chiama retta.

In secondo luogo si dovevano ancora essere sistemate le operazioni fondamentali con gli infinitesimi. Questi complicano ulteriormente il quadro del sistema numerico, perchè dovrebbero rappresentare una nuvoletta di numeri molto vicini a zero, minori di ogni numero per quanto piccolo. Non si vedono, ma se si avesse un microscopio trascendente si vedrebbe questa nuvoletta, o in omaggio a Leibniz una *monade*, e non solo intorno a zero, ma per trasla-

[15] J. Dieudonné, *Abrégé* ..., vol. I, cit., p. 326.

zione intorno a ogni numero reale. Gli infinitesimi erano utili e usati, ma la decisione è stata quella di eliminarli.

Il primo passo per l'eliminazione degli infinitesimi è attribuito di solito a Cauchy (ma predecessori ce ne sono sempre, ad esempio d'Alembert, e certe formulazioni attente dello stesso Newton) che propone la definizione di limite corretta e attuale, benché ancora formulata in linguaggio naturale, con gli avverbi "arbitrariamente" e "definitivamente", e non accompagnata in modo coerente da una chiara nozione delle quantità variabili. La definizione che si è imposta con ϵ e δ si deve soprattutto a Weierstrass, a cui si deve anche la prima definizione aritmetica del continuo, verso la metà del secolo, superata poi da quelle più soddisfacenti di Cantor e di Dedekind (1872).

Dedekind si ispira per la sua definizione alla caratteristica della continuità rappresentata dall'assenza di lacune; i suoi numeri reali sono le partizioni dei razionali in due *sezioni* superiore e inferiore, in mezzo alle quali deve sempre esservi un numero; la proprietà di continuità o completezza (si veda oltre), espressa dal cosiddetto teorema di Bolzano-Weierstrass, che ogni insieme limitato superiormente ha un estremo superiore, risulta in questo contesto di facile dimostrazione, più che non la definizione stessa delle operazioni algebriche.

Nella versione di Cantor, i reali siano invece successioni di razionali. Entrambe richiedono una maggior disinvoltura e sicurezza nel maneggiare l'infinito attuale, e nel dominare un linguaggio sempre più ricco di termini insiemistici, progresso che era in corso grazie in particolare proprio al lavoro di Cantor e di Dedekind.

Le definizioni del continuo si dicono aritmetiche perché si basano essenzialmente sui numeri razionali, con un modico di contributi insiemistici che non si distinguono tuttavia chiaramente dalla logica o meglio dal linguaggio naturale (il termine "insieme" è usato inizialmente, si potrebbe dire, come una delle nozioni comuni euclidee). I numeri razionali a loro volta, come gli interi relativi, si definiscono a partire dai numeri naturali, di nuovo con un modico di insiemistica.

Tutto l'edificio dei sistemi numerici poggia infine in modo soddisfacente, e coerente con l'intuizione aritmetica e algebrica, sui numeri naturali. Resta da sapere cosa sono i numeri naturali, ché sembra di essere ancora nella situazione in cui ci si trovava al tempo del primo sviluppo euclideo dell'aritmetica.

Anche questa lacuna viene finalmente colmata e si trovano gli assiomi per i numeri naturali, di nuovo con una scoperta plurima.

Merita in particolare di essere conosciuto il percorso di Dedekind, sostanzialmente equivalente alla elaborazione di Frege, ma più facile da esporre in una terminologia familiare. Dedekind definisce la struttura dei naturali come il più piccolo insieme infinito.

La definizione di "infinito" di Dedekind è diretta, invece di assumere la usuale connotazione negativa di "infinito" come "non finito".

Un insieme è infinito secondo Dedekind se esiste un'iniezione propria dell'insieme in sé.

Se si richiede per un insieme infinito anche la minimalità, ne viene che un solo elemento, chiamiamolo 0, è l'unico elemento che non appartiene all'immagine dell'iniezione - funzione che ora si chiama "il successore di" - e ne viene anche la validità del principio di induzione e la conseguente possibilità delle definizioni per ricorsione. Sono questi essenzialmente gli assiomi dei numeri naturali.

Il ragionamento seguito da Giuseppe Peano per arrivare agli stessi assiomi di Dedekind non è noto, perchè non è stato da lui dichiarato; Peano era molto attento e abile a cogliere le proposizioni e le leggi ricorrenti nelle dimostrazioni e di uso più frequente e fondamentale; grazie a questa sua capacità poté ad esempio assiomatizzare la nozione di spazio vettoriale. Peano conosceva Grassmann e aveva probabilmente recepito dalla sua opera l'importanza delle definizioni per ricorsione, che egli utilizza sistematicamente per le operazioni aritmetiche.

Peano era un assiomatizzatore, e quindi presenta i suoi assiomi proprio come tali; Dedekind colloca la sua analisi e le sue definizioni nell'ambito di una trattazione degli insiemi (che egli chiama sistemi).

Nonostante l'impostazione non assiomatica tuttavia, alla fine Dedekind riassume la sua definizione dicendo che "l'essenza di un insieme semplicemente infinito $\mathbb{N}$" è data da alcune proprietà, che elenca, e che sono le stesse di Peano, o forse più precise, perché Peano non postula l'iniettività del successore (per lui "+1") ma solo il suo carattere funzionale; quindi che 1 non è un successore[16] e l'induzione.

L'apparente coincidenza dei due procedimenti sarà nel seguito fonte di confusione, unitamente al fatto che Peano usa un linguaggio con classi[17].

[16] Peano parte da 1, non da 0.

[17] Quelli che ora si chiamano assiomi di Peano non sono quelli originali. Non si usa il linguaggio insiemistico, e l'induzione è formulata come schema; per dimostrare il teorema di ricorsione occorre avere almeno gli insiemi finiti; una codifica di questi con i numeri è possibile se si hanno a disposizione l'addizione e la moltiplicazione; le definizioni ricorsive

2.9 Hilbert e la geometria

Il lavoro di Hilbert sulla geometria è molto importante per capire come le preoccupazioni metamatematiche siano diventate parte della matematica.

Gli obiettivi che Hilbert si poneva quando iniziò a dedicarsi allo studio della geometria, che sarebbe culminato nelle *Grundlagen*, sono stati da lui stesso brevemente descritti nell'epistolario con Gottlob Frege[18].

"Io sono stato costretto dalla necessità a stabilire il mio sistema di assiomi: volevo rendere possibile la comprensione di quelle, fra le proposizioni della geometria, che ritengo essere i risultati più importanti delle indagini geometriche: ossia che l'assioma delle parallele non è conseguenza dei rimanenti assiomi, che lo stesso vale per l'assioma di Archimede ecc. Volevo rispondere alla domanda tendente a stabilire se la proposizione che in due rettangoli equivalenti e di ugual base sono uguali anche gli altri lati, può venir dimostrata o vada piuttosto assunta come un nuovo postulato, come già fa Euclide[19]. In generale, volevo stabilire la possibilità di comprendere e dare una risposta a domande del tipo: perché la somma degli angoli interni

---

di queste sono allora aggiunte agli assiomi.

[18]Il carteggio tra Hilbert e Frege è pubblicato in G. Frege, *Alle origini della logica moderna*, Boringhieri, Torino, 1983. Esso si svolse tra il 1901 e il 1905, e vi torneremo in **3.2**.

[19]Hilbert potrebbe riferirsi qui a quanto osservato nelle *Grundlagen* a proposito del teorema 48, che afferma che due triangoli equicomplementabili con la stessa base hanno anche la stessa altezza. Secondo Hilbert questo teorema si trova in Euclide I. 39, e "nella dimostrazione Euclide fa appello al teorema generale sulle grandezze che il tutto è maggiore di una sua parte, un metodo che è equivalente alla introduzione di un nuovo assioma geometrico di equicomplementabilità". Che il tutto sia maggiore delle sue parti non è un teorema ma una nozione comune, per Euclide. Hilbert pensa che il suo uso nel contesto abbia un contenuto geometrico, da denunciare come tale. Infatti considera nell'Appendice II una geometria con un indebolimento del suo assioma III. 5 di congruenza (il caso di due lati e angolo compreso), nella quale non vale il teorema 48, e quindi neanche il principio sul tutto e le parti.

La proposizione I. 39 non è letteralmente uguale a quella citata da Hilbert, ma afferma che triangoli uguali che abbiano la stessa base sono nelle stesse parallele, cioè la retta che unisce i due vertici è parallela alla base (e compare il riferimento alla nozione comune). Può darsi che Hilbert la considerasse la più vicina a quelle che gli servono per il suo teorema.

Euclide usa lo stesso metodo in altri teoremi su triangoli e congruenze che sono imparentati con il teorema 48 di Hilbert, ad esempio per I. 26, la proposizione che afferma che se due triangoli hanno due angoli e un lato uguali, il lato essendo o quello che unisce i due angoli uguali o uno che sottende uno dei due angoli uguali, allora hanno uguali anche gli altri lati e angoli.

di un triangolo vale due retti? E volevo inoltre chiarire come questo fatto fosse collegato all'assioma delle parallele".

Ma a parte questo interesse per la geometria piana, il contenuto dell'opera di Hilbert non si apprezza se non si tiene conto di altre sue preoccupazioni, forse troppo complicate da esporre in una lettera. Il contesto in cui si collocano le ricerche di Hilbert è quello in cui la geometria, come abbiamo accennato, era stata sbalzata dal suo ruolo di regina della matematica, a favore della teoria dei numeri.

Mentre nell'Ottocento l'Analisi cresceva rigogliosa, la Geometria era piuttosto rivolta su se stessa ad interrogarsi sulla propria funzione e sul proprio significato; il ruolo dell'intuizione geometrica era scaduto a livelli minimi di affidabilità; contribuiva alla sua svalutazione lo studio delle funzioni non più rappresentabili da curve e con patologie dominabili solo con la nuova intuizione insiemistica dell'infinito.

La geometria rientrava nella matematica pura solo grazie all'analitica, quindi la sua legittimità e giustificazione dipendevano da quelle dei numeri. La geometria in sé trattava grandezze continue, e per questo non era considerata matematica pura, né da Gauss né da Kronecker. Era questo il periodo dell'aritmetizzazione, quando Kronecker sosteneva che i numeri naturali sono dati da Dio, il resto è opera dell'uomo.

Sulla dichiarazione di Kronecker spesso si sorvola, la si prende come una battuta curiosa, da citare senza sapere bene cosa voglia dire - forse la natura divina della matematica - comunque una sua idea personale; invece essa è molto significativa, perché rivela la completa rinuncia ad una spiegazione della matematica: non richiedeva una definizione del numero, che ancora non era stata data, ma non si preoccupava neanche di trovarne gli assiomi, tanto è vero che li affidava a Dio.

La prima ambizione di Hilbert è stata quella di (ri)dare alla geometria uno *status* e una legittimazione autonoma, indipendente dal numero.

Un secondo obiettivo è stato quello di chiarire il ruolo degli assiomi di continuità in geometria. La continuità era al centro della preoccupazione sia per quel che riguardava i fondamenti dell'Analisi sia per quel che riguardava la natura dello spazio e la decisione di come dovesse essere rappresentato. Cantor aveva mostrato che (la definizione de) il moto continuo era compatibile con uno spazio totalmente sconnesso. La geometria non pareva più in grado di fornire lei, nonostante si dicesse che trattava grandezze continue, un fondamento della continuità, che veniva ricercato nelle definizioni aritmetiche del continuo.

Infine tra i problemi interni alla teoria geometrica in quanto tale spiccavano in modo particolare quelli che riguardavano i rapporti tra le varie dimensioni.

Euclide non aveva trattato la geometria dello spazio, salvo per alcuni solidi. Negli *Elementi* la geometria piana non fa mai uso dello spazio, a differenza delle costruzioni dei geometri che trattavano i problemi classici. Un problema rilevante a questo proposito e molto dibattuto è quello del teorema di Talete su triangoli isosceli, dove viene usata una riflessione, che sembra richiedere di uscire nello spazio. Anche Proclo ammette che lo spazio ha una natura controversa e difficile da scoprire.

Nei secoli di studio della geometria si erano aggiunti agli assiomi di Euclide altri assiomi mancanti, precisamente assiomi per l'ordine, la congruenza, la continuità, lo spazio.

Quando nella seconda metà dell'Ottocento, anche dopo le geometrie non euclidee, si contemplavano le aggiunte che erano state fatte o si dovevano fare a Euclide, poteva sorgere il dubbio che il quinto postulato potesse diventare dimostrabile con i nuovi assiomi, in particolare con quelli dello spazio (fino ad allora non molto esplorati in rapporto alle parallele).

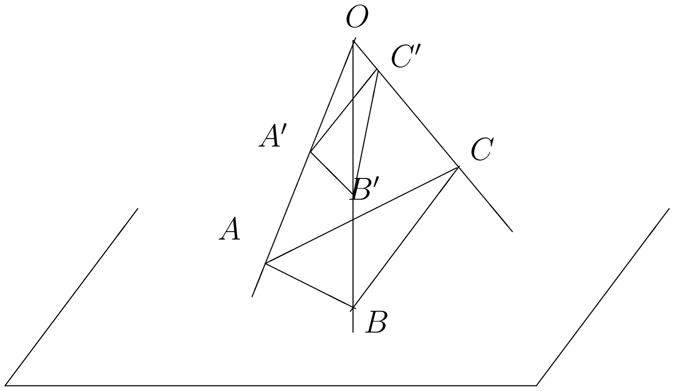
Bolyai ad esempio era tra quelli convinti di ciò, perchè osservava che molte proprietà del piano diventano evidenti nel contesto dello spazio.

Chi aveva esplicitamente trattato lo spazio, cercando di rappresentare la terza dimensione nel piano, era stato naturalmente Desargues con la geometria proiettiva.

L'ambizione di Desargues era quella di dare una trattazione teorica della prospettiva, ed egli aveva presentato anche un abbozzo assiomatico. Definiva i punti all'infinito come il fine dove rette parallele tendono a incontrarsi. Due triangoli ABC e $A'B'C'$ erano detti "prospettici" da O se le rette che univano i lati omoghi AA', BB', CC' si incontravano in O. Mentre i triangoli simili sono quelli che hanno la stessa forma, i triangoli prospettici (da O) sono quelli che hanno da O la stessa apparenza.

Il cosiddetto teorema di Desargues dà un criterio perché due triangoli siano prospettici. Il teorema di Desargues ha avuto un ruolo fondamentale nella riflessione sulla geometria nel corso dell'Ottocento, non solo in Hilbert, ma anche in studiosi come C. S. Peirce, per non parlare di quelli come von Staudt che ne diedero nuove dimostrazioni.

Il teorema afferma che se due triangoli ABC e $A'B'C'$ sono prospettici, i tre punti in cui si incontrano le rette che uniscono i lati corrispondenti AB e $A'B'$, AC e $A'C'$, BC e $B'C'$ sono allineati, come si può vedere completando la figura



Come proposizione su triangoli nello spazio il teorema è immediato, utilizzando solo gli assiomi di incidenza di piani; come proposizione su triangoli in un piano invece è più difficile. La dimostrazione è ancora abbastanza agevole se si esce fuori dal piano, nello spazio, proiettando tutta la configurazione da un punto esterno; allora bastano di nuovo gli assiomi di incidenza. Ma se si vuole dimostrarlo come teorema della geometria piana, la dimostrazione è assai complicata. Desargues lo dimostrò usando congruenza (il criterio dei due lati e angolo compreso, ma applicato a due triangoli di cui uno è riflessione dell'altro, che quindi in un certo senso richiedono lo spazio), similarità e continuità[20].

Hilbert dedicò molta attenzione al teorema di Desargues. Egli dimostrò che il teorema piano non era derivabile (nella geometria piana) neanche con ordine, parallele e continuità; è necessario l'assioma della congruenza dei triangoli, come nella dimostrazione originale. Lo spazio non era dunque

[20]S'intende: assiomi di congruenza ecc.

evitabile a meno di non ricorrere ad altre nozioni - congruenza - che erano estranee alla formulazione dell'enunciato del teorema.

Questo fu l'inizio di una costante preoccupazione di Hilbert per la cosiddetta *purezza dei metodi* (che si ottiene quando in un dominio è possibile evitare il ricorso a nozioni e strumenti più potenti, o anche semplicemente diversi).

Tale problematica si collega a quella della completezza, di cui diremo, ed è anche nel suo programma di fondazione della matematica.

Ma l'analisi di Hilbert andò oltre: egli dimostrò che un piano proiettivo (cioè un piano in cui valgano gli assiomi di incidenza e ordine) può essere immerso nello spazio se e solo se in esso vale l'enunciato del teorema di Desargues.

L'enunciato di Desargues si comporta come una nuova regola per la rappresentazione dello spazio nel piano, ed essa si è rivelata fondamentale; è come se fornisse agli abitanti di Flatlandia un modo di intuire lo spazio, molto meglio dell'artificio inventato da Abbott.

Nella dimostrazione del teorema di immersione di Hilbert si deve definire uno spazio in cui immergere il piano, e i punti dello spazio sono triangoli del piano originario[21] e le rette terne di rette. Questa soluzione è ben più significativa, quanto alla natura non fissata degli enti geometrici, che non il pensarli come amore, legge o spazzacamini[22].

Infine Hilbert usa Desargues e la definizione cartesiana di moltiplicazione (senza congruenza né Archimede) per costruire un corpo. Senza l'assioma di Archimede il corpo risulta non commutativo; l'assioma di Archimede serve per dimostrare la commutatività[23]. (Dell'assioma di completezza diremo dopo.)

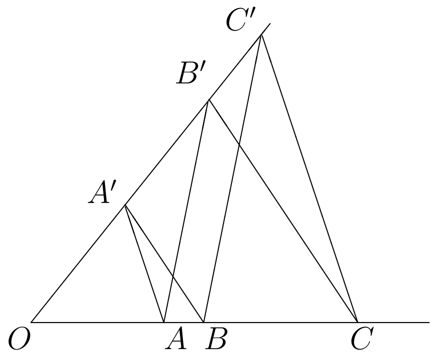
Nelle *Grundlagen* si trovano altri risultati, sempre orientati a chiarire il ruolo dei vari assiomi, ad esempio sui rapporti tra il teorema di Desargues e quello di Pascal, che afferma che per sei punti qualunque, tre su una retta e tre su un'altra, se AB' è paralela a BC' e AA' è parallela a CC' allora BA' è parallela a CB'

---

[21] Si fissa un triangolo e si considerano tutti i triangoli che hanno i lati paralleli a quelli del triangolo fissato.

[22] Una battuta di Hilbert, a proposito della definizione degli enti geometrici, su cui torneremo.

[23] Oppure invece di Desargues e Archimede si usa il teorema di Pascal, che implica Desargues e non ne è implicato senza la continuità, si veda oltre.



Il teorema non esprime altro che la commutatività della moltiplicazione. Tutti questi risultati non sono tuttavia direttamente pertinenti all'attuale discussione.

Con il complesso delle ricerche presentate nelle *Grundlagen*, Hilbert realizza la sua ambizione originaria: nella geometria analitica gli enti geometrici dipendono dai numeri; ma rovesciando l'analitica, partendo dalla geometria, si ottiene un corpo delle stesse caratteristiche dei reali. Numeri ed enti geometrici non sono tanto diversi, entrambi soddisfano gli assiomi dei corpi[24].

[24]In particolare si ottiene in tal modo un corpo reale chiuso, un modello di una teoria che Tarski dimostrerà in seguito completa e decidibile (si veda più avanti per le definizioni).

3 Non contraddittorietà e completezza

Ritorniamo a considerare i nuovi sistemi assiomatici di fine Ottocento. Dalla loro proliferazione emerge un problema metalogico, quella della loro non contraddittorietà, o compatibilità, o coerenza (ingl. *consistency*, qualche volta italianizzato in "consistenza").

Una teoria T - ovvero un insieme di assiomi - si dice non contraddittoria se non è possibile dimostrare in essa una contraddizione, ovvero, con le notazioni della logica, non esiste un enunciato φ tale che[1]

$$T \vdash \varphi \wedge \neg\varphi$$

oppure (in modo equivalente) tale che separatamente $T \vdash \varphi$ e $T \vdash \neg\varphi$.

3.1 Non contraddittorietà

La compatibilità di un sistema di assiomi è un requisito che si impone tecnicamente quando si voglia dimostrare una indipendenza, cioè la compatibilità della negazione. Infatti

$$T \cup \{\varphi\} \text{ è non contraddittorio se e solo se } T \nvdash \neg\varphi.$$

Ma la preoccupazione della coerenza, per le nuove teorie come per le classiche, era anche sintomo di una incertezza reale.

Pasch non aveva incertezze sulla geometria che studiava. Per lui "i concetti geometrici sono certo corrisposti, in origine, a degli oggetti empirici", anche se poi "si sono chiusi a poco a poco in una rete di concetti artificiali al fine di favorire lo sviluppo teorico"; i concetti di origine empirica sono i concetti primitivi, mentre quelli che si sono aggiunti sono i derivati. I concetti primitivi secondo Pasch è inutile cercare di spiegarli; nessuna spiegazione sarebbe meglio che il rinvio agli appropriati oggetti della natura a cui corrispondono nella loro origine.

[1] $\vdash$ è il segno di derivabilità, $\neg$ il segno di negazione e $\wedge$ il segno di congiunzione; $\varphi\wedge\neg\varphi$ si dice una contraddizione. $T \vdash \varphi\wedge\neg\varphi$ si legge "$\varphi\wedge\neg\varphi$ è derivabile da T" oppure "da T si deriva, deduce $\varphi \wedge \neg\varphi$". $\nvdash$ significa "non è derivabile". Veramente il segno $\vdash$ si dovrebbe riferire a una logica precisata, e alle sue regole, ma lo usiamo per indicare una logica qualunque, anche quella che si usa informalmente nelle dimostrazioni, perché le regole che permettono queste e le prossime definizioni preliminari valgono per tutte.

Analogamente per le dimostrazioni, per Pasch i principi primi sono quelli che sono fondati in maniera immediata su osservazioni, che da tempo immemorabile si sono ripetute.

Invece Poincaré, dopo aver dichiarato che gli assiomi sono convenzioni, si chiedeva: "è certo che tutte queste convenzioni sono compatibili?

Queste convenzioni, è vero, ci sono state suggerite dalle esperienze, ma da esperienze grossolane. Noi scopriamo che si verificano approssimativamente certe leggi e decomponiamo per convenzione il fenomeno osservato in due altri: un fenomeno puramente geometrico che obbedisce esattamente a quelle leggi e un piccolo fenomeno perturbativo. È certo che questa decomposizione sia sempre legittima? ..."[2].

Le varie geometrie avevano rinunciato a essere fondate sulla realtà, ma almeno il requisito logico della non contraddittorietà sembrava necessario. D'altra parte la definizione stessa di "teoria matematica" a cui si era pervenuti coscientemente ed esplicitamente richiedeva per tutte la garanzia della non contraddittorietà; da una contraddizione segue qualunque cosa, per la legge *ex falso quodlibet*

$$\varphi \wedge \neg\varphi \to \psi$$

per ogni φ e ψ, e quindi una teoria contraddittoria è banale, tutto il lavoro fatto per svilupparla un *love's labour lost*[3].

La riduzione di tutta la matematica ai numeri naturali legava tra di loro le diverse teorie classiche anche per quel che riguarda la non contraddittorietà.

Le diverse geometrie euclidea e non euclidee erano relativamente mutuamente non contraddittorie: i modelli di geometrie non euclidee in figure definite nella geometria euclidea (come la sfera e la pseudosfera) permettono una giustificazione condizionale di una rispetto all'altra.

Si ha in particolare un rinvio dall'una all'altra della proprietà di non contraddittorietà. Si definisce ad esempio una geometria sulla superfice sferica

[2]H. Poincaré, "On the Foundations of Geometry", cit.

[3]Esistono filosofi che pensano che una contraddizione non sia globalmente distruttiva, ma possa avere un effetto locale senza influenzare altre parti della teoria o deprezzare i teoremi in essa dimostrati. Per giustificare una simile posizione occorre come minimo pensare a una logica soggiacente senza *ex falso quodlibet*, e alcune sono state sperimentate. Naturalmente quando si scopre una contraddizione in una teoria fino ad allora ritenuta coerente, il lavoro fatto, nel senso delle dimostrazioni corrette costruite, può essere recuperato in un quadro corretto; è successo con il sistema logico di Frege, dimostrato contraddittorio dall'antinomia di Russell.

ponendo che i "punti" siano le coppie di punti opposti della superficie, individuati dall'intersezione con la superficie di una retta per il centro (diametro), e le "rette" siano i cerchi massimi. I primi assiomi sono soddisfatti, ad esempio per due "punti" passa una "retta": dati due "punti" infatti, questi con i diametri loro associati individuano un piano passante per il centro la cui intersezione con la superfice sferica è un cerchio massimo. Ed è facile costruire un triangolo con tre angoli retti; si prenda l'equatore e due meridiani per i poli perpendicolari tra loro. Attraverso la corrispondenza

$$\begin{gathered}\text{coppie di punti opposti} \rightsquigarrow \text{punti}\\ \text{cerchi massimi} \rightsquigarrow \text{rette}\end{gathered}$$

si ha un'interpretazione che stabilisce la *non contraddittorietà relativa* della geometria non euclidea ellittica rispetto a quella euclidea[4].

Una teoria T_1 si dice non contraddittoria relativamente a una teoria T_2 se la non contraddittorietà di T_2 implica la non contraddittorietà di T_1.

Se si avesse una contraddizione nelle proprietà di questa nuova geometria, una contraddizione riguardante le "rette" ad esempio, questa sarebbe una contraddizione riguardante i cerchi massimi di una sfera euclidea, e le geometria euclidea sarebbe contraddittoria.

Viceversa, la contraddittorietà della geometria euclidea avrebbe come conseguenza la possibilità di dimostrare qualunque cosa, anche contraddizioni relative ai cerchi massimi di una sfera, e quindi una contraddizione nella geometria non euclidea.

Ma attraverso la geometria analitica la non contraddittorietà della geometria era ricondotta a quella dei numeri reali, e questa in ultima analisi a quella dei numeri naturali, perché nessun dubbio sussisteva sulla correttezza e validità dei metodi usati per definire i sistemi numerici: per gli interi e i razionali si trattava in effetti solo di formazione di classi di equivalenza definibili; lo stesso sembrava per i reali, anche se di fatto occorreva qualcosa di più, che non era tuttavia chiaramente individuato[5].

[4]In questo caso l'interpretazione non è una struttura ma consiste nella sostituzione di concetti, o predicati, a concetti: le formule di una teoria sono trasformate in formule di un'altra teoria con una trasformazione effettiva, che permette di trasportare in modo effettivo una contraddizione eventuale dall'una all'altra. Corrisponde a quello che veniva chiamato cambiamento di significato. La sfera con le sue "rette" e "punti" sarebbe invece una struttura, che si chiama anche *modello* della teoria.

[5]In una ottica assiomatica, occorre l'uso pieno dell'assioma della potenza, per insiemi infiniti.

Non che si temesse davvero la contraddittorietà, dopo millenni di esperienza; ma gli assiomi dell'aritmetica dipendevano dalla nozione di infinito, e nella teoria degli insiemi *in fieri* contraddizioni se ne erano trovate, e l'Analisi usava pesantemente ragionamenti pericolosamente vicini a quelli che portavano a contraddizioni[6].

Come si dimostra la compatibilità di uno dei sistemi di assiomi della matematica? e in particolare di quello dell'aritmetica? S'intende una non contraddittorietà assoluta, non relativa ad ancora un'altra teoria. Enriques delinea brevemente due tendenze. "Alcuni ritengono che la compatibilità di un sistema d'ipotesi possa provarsi soltanto a posteriori, in quanto si assegni al sistema una interpretazione qualsiasi nel dominio dell'esperienza. Gli altri (sull'esempio di Weierstrass, Kronecker ... ecc.) tendono a ridurre il significato astratto del sistema all'interpretazione aritmetica, dove tutto appare fondato sul concetto dei numeri naturali: del quale si ammette a priori la possibilità, in base alle leggi del pensiero". Questi altri in sostanza rinunciano, o non ritengono necessaria, una dimostrazione della non contraddittorietà dell'aritmetica. Una terza posizione veramente era quella del programma di Hilbert, che Enriques non affronta nemmeno[7], cioè di provare a dimostrare la non contraddittorietà dell'aritmetica, dopo averla formalizzata e ridotta (teoremi e dimostrazioni) a un insieme di strutture simboliche finite, usando strumenti matematici di estrema affidabilità, una matematica che si riducesse a manipolazioni di oggetti concreti e fosse quindi una pre-condizione della matematica stessa.

Enriques tende a concordare con i secondi, accettando per quel che riguarda i numeri naturali che si tratti di una base ultima; egli cita con approvazione la posizione di Poincaré, secondo il quale il principio di induzione è un principio sintetico di esistenza, a cui si riconducono tutte le altre dottrine matematiche[8]

3.2 Amore, leggi, spazzacamini

Un altro problema molto più complesso che emerge è quello della completezza, collegato a quello della molteplicità delle interpretazioni. Per spiegarlo, è

[6] Ad esempio le cosiddette definizioni impredicative nel teorema di Bolzano-Weierstrass.

[7] Enriques scrive nel 1922; Hilbert aveva accennato alla sua idea nel 1904, ma solo nel 1917 era tornato a dedicarsi a studi fondazionali e a proporre il suo programma.

[8] Per parte sua Enriques svolge un'analisi mentale del processo logico, che porta secondo lui ad una esistenza mentale, dove sarebbero escluse le contraddizioni.

opportuno riassumere di nuovo le caratteristiche del metodo ipotetico deduttivo, questa volta con le parole di Hilbert, più precise e consapevoli di quelle di Enriques.

Hilbert non aveva concepito il suo lavoro delle *Grundlagen* come un manifesto del nuovo metodo assiomatico, o come un esercizio logico di assiomatizzazione; abbiamo visto da quali interessi matematici fosse mosso. Se aveva le sue idee, nel 1899 aveva preferito lasciar parlare la nuda esposizione. Ma alla fine è costretto a pronunciarsi essendo tirato per la manica dal logico Gottlob Frege.

Un sistema di assiomi non definisce un concetto, diceva Frege, perchè manca alla definizione il requisito della dimostrazione di esistenza. "Il sistema di definizioni che Lei usa è paragonabile a un sistema di equazioni a più incognite per il quale rimane dubbia la risolubilità, e in particolare l'univocità della determinazione delle incognite". Quello che disturbava Frege e molti matematici non era tanto la mancanza della dimostrazione di esistenza per la definizione assiomatica, quanto la pluralità delle interpretazioni. Non Enriques come abbiamo visto, ma il problema diventa scottante per l'aritmetica se si vuole assicurare uno stato speciale ai numeri naturali, tra i possibili diversi modelli; per i numeri naturali Enriques si appellava alla natura dei giudizi sintetici.

Un buon sistema di assiomi, uno che esprimesse l'essenza del concetto, raggiunto attraverso un'analisi corretta ed esauriente della nozione intuitiva, doveva nel pensiero tradizionale individuare una sola realtà, il *genus*, come in una definizione reale, e quindi doveva essere tale da avere una sola interpretazione.

Essendo questi i suoi presupposti, Frege non si orienta nelle *Grundlagen*:

> Nel paragrafo 6 Lei dice: "Gli assiomi di questo gruppo definiscono il concetto della congruenza o del movimento". Ma allora, perché mai essi non vengono chiamati definizioni? ... A tutta prima vien fatto di pensare ai punti nel senso della geometria euclidea, e la Sua affermazione - che gli assiomi esprimono fatti fondamentali della nostra intuizione - conferma tale opinione. In seguito però Lei intende per punto una coppia di numeri. Resto dubbioso di fronte alle affermazioni che per mezzo degli assiomi della geometria si raggiunge la descrizione completa e precisa delle relazioni, e che gli assiomi definiscono il concetto del "fra". Con ciò si ascrive agli assiomi qualcosa che è compito delle definizioni.

> Così facendo vengono - a mio parere - seriamente confusi i confini tra assiomi e definizioni, e accanto al significato tradizionale della parola "assioma" - quale risulta nell'affermazione che gli assiomi esprimono fatti fondamentali dell'intuizione - mi sembra ne affiori un secondo, che peraltro non mi riesce di cogliere esattamente ...
>
> Attribuisco il nome di assiomi a enunciati che sono veri, ma che non vengono dimostrati perché la loro conoscenza scaturisce da una fonte conoscitiva di natura extra-logica, che possiamo chiamare intuizione spaziale. Il fatto che gli assiomi siano veri ci assicura di per sé che essi non si contraddicono tra loro, e ciò non abbisogna di alcuna ulteriore dimostrazione[9].

"Il punto cardinale dell'equivoco" di Frege per Hilbert si riassumeva nella seguente spiegazione:

> Io non voglio presupporre nulla come noto; io vedo nella mia spiegazione del paragrafo 1 la definizione dei concetti di punto, retta, piano, se si tornano ad assumere come note caratteristiche tutti gli assiomi dei gruppi I-V. Se si cercano altre definizioni di "punto", ricorrendo per esempio a perifrasi come "privo di estensione" ecc., si capisce che debbo oppormi nel modo più deciso a siffatti tentativi; si va infatti alla ricerca di qualcosa là dove non la si potrà mai trovare, per il semplice motivo che non è là dove la si cerca.

Quindi in riferimento all'affermazione di Frege che gli assiomi sono veri osservava:

> Mi ha molto interessato leggere nella Sua lettera proprio questa frase, poiché io, da quando ho cominciato a riflettere, scrivere e tenere conferenze su questo argomento, ho sempre detto esattamente il contrario: se assiomi arbitrariamente stabiliti non sono in contraddizione, con tutte le loro conseguenze, allora essi sono veri, allora esistono gli enti definiti per mezzo di quegli assiomi. Questo è per me il criterio della verità e dell'esistenza ... Lei dice ... che ad esempio "fra" è concepito in modo diverso a pagina

[9]Anche gli italiani Peano e Padoa erano convinti di ciò, in relazione all'aritmetica: la prova della non contraddittorietà degli assiomi dell'aritmetica è $\mathbb{N}$.

> 20 e che ivi il punto è una coppia di numeri. Certamente, si comprende da sé che ogni teoria è solo un telaio, uno schema di concetti unitamente alle loro mutue relazioni necessarie, e che gli elementi fondamentali possono venir pensati in modo arbitrario. Se con i miei punti voglio intendere un qualunque sistema di enti, per esempio il sistema: amore, legge, spazzacamino ..., allora basterà che assuma tutti i miei assiomi come relazioni tra questi enti perché le mie proposizioni, per esempio il teorema di Pitagora, valgano anche per essi. In altre parole: ogni teoria può essere sempre applicata a infiniti sistemi di elementi fondamentali. Anzi occorre soltanto applicare una trasformazione biunivoca e convenire che gli assiomi per gli enti trasformati debbano essere uguali a quelli che valgono per i loro corrispondenti. Di fatto anche questa circostanza si applica sovente, ad esempio col principio di dualità ecc., e io l'applico alle mie dimostrazioni di indipendenza.

Il "criterio della verità e dell'esistenza" di Hilbert mette in luce un altro motivo per la richiesta della non contraddittorietà di una teoria: la non contraddittorietà implica l'esistenza. Si spiega così l'interesse esasperato di Hilbert - e della maggioranza dei matematici del tempo, che lo hanno seguito - per la non contraddittorietà dell'aritmetica.

Bisogna notare tuttavia che il criterio di Hilbert non è una sua credenza filosofica, o non è soltanto questo, anche se non sappiamo come vi sia arrivato: è l'intuizione e l'anticipazione di una verità che sarà *dimostrata* da Kurt Gödel nel 1929 con il teorema di completezza logica, come vedremo.

Un corollario di questo teorema, o meglio una versione equivalente, afferma precisamente che

se T è non contraddittoria, allora esiste un'interpretazione in cui T è vera,

dove con "interpretazione" si intende proprio una struttura nella quale tutti gli assiomi di T sono veri.

Alla fine non riuscendo a convincere l'interlocutore Hilbert chiude la discussione: "La mia idea è appunto questa: un concetto può essere logicamente definito solo attraverso le sue relazioni con altri concetti. Queste relazioni, formulate in enunciati determinati, le chiamo assiomi e pervengo in tal modo al risultato che gli assiomi ... sono le definizioni dei concetti. Non ho maturato questa concezione come scorciatoia, ma al contrario vi sono stato costretto dalla ricerca del rigore nella deduzione logica e nella costruzione logica di una teoria".

3.3 Completezza

Se il metodo assiomatico sembrava evadere l'esigenza della unicità della interpretazione, una scappatoia c'era, quella di (provare a) provare, certamente nel caso dei numeri naturali, ma anche in quelli dei reali e della geometria, che di interpretazioni ce ne era moralmente una sola. Se ce ne erano diverse - perché c'erano, come minimo prendendo sistemi in corrispondenza biunivoca con uno dato, come dice Hilbert - le differenze dovevano essere inessenziali. I diversi sistemi dovevano essere appunto solo quelli che si ottenevano con corrispodenze biunivoche. Diversi sistemi numerici (in corrispondenza biunivoca) potevano differire per caratteristiche non interessanti, non matematiche, come se i loro elementi avessero un colore diverso. Quindi dovevano essere, o si sperava che fossero *isomorfi*.

La proprietà di una teoria di avere i modelli tutti isomorfi tra loro venne chiamata in seguito *categoricità*.

L'insiemista Adolf Fraenkel diceva ancora nel 1928[10] che nel caso di un sistema di assiomi categorico si ha una determinazione formalmente univoca, o completa dei concetti primitivi, e quindi un sostituto formalmente valido di una definizione, salvando tuttavia la pluralità delle interpretazioni del metodo assiomatico e la possibilità di sfruttare quelle interpretazioni che fossero più intuitive:

> Quando per un determinato significato concreto (*inhalthichen*) dei concetti primitivi (per esempio "punto" e "retta" nel senso intuitivo) una proposizione risulta valida (*richtig*), cioè deduttivamente derivabile dagli assiomi, allora la proposizione non può essere falsa rispetto a un altro significato pure compatibile con gli assiomi (per esempio "punto" come "coppia di numeri"), altrimenti si avrebbe una contraddizione con l'isomorfismo dimostrato. E tuttavia questo non significa che il senso, il contenuto essenziale dei concetti primitivi possa mai essere determinato dagli assiomi, perché con ogni interpretazione ce ne è una sia pure isomorfa ma il cui senso è un altro.

[10] A. A. Fraenkel, *Einleitung in die Mengenlehre*, Springer, Berlin, 3^a ed., 1928.

All'inizio del secolo tuttavia la terminologia non è sedimentata, anche perché i concetti non sono del tutto chiari. Il quadro si arricchisce prendendo in considerazione la scuola americana di assiomatizzatori. Nel 1902 in un lavoro sulle grandezze continue Huntington propone un insieme di sei assiomi che afferma essere *completo*, intendendo con questo che gli assiomi sono "compatibili, sufficienti e mutuamente indipendenti". "Compatibili" significa che "esiste almeno un insieme (*assemblage*) nel quale le regole di combinazione soddisfano tutti e sei i requisiti", mentre "sufficienti" significa che "esiste essenzialmente *un solo* insieme del genere"[11].

Anche Veblen nel 1904, a proposito di un suo sistema di geometria si esprime affermando che "esiste *essenzialmente una sola* classe nella quale gli assiomi sono validi", e usa per questa proprietà del sistema il termine "categorico", attribuendolo al filosofo John Dewey, che per il caso multiplo avrebbe suggerito il termine "disgiuntivo"[12], non fortunato come il primo e non impostosi. Ma il termine "completo" resta per alcuni anni con il significato di "categorico". Lo usa così ancora Enriques nel 1922, che non era aggiornato sugli sviluppi della logica: un sistema è completo se "due sistemi di enti costretti a soddisfare al sistema possano porsi tra loro in corrispondenza biunivoca, in siffatta guisa che le proprietà dell'uno si traducano in proprietà perfettamente omologhe dell'altra, sicché appariscano astrattamente *uguali*, nell'ordine d'idee di cui si tratta"[13].

Hilbert usa pure il termine "completezza" (*Vollständigkeit*), in un'accezione che sembra simile alla categoricità, ma con una motivazione più complessa e che nasconde altre preoccupazioni.

All'inizio delle *Grundlagen* Hilbert dichiara che il suo obiettivo è quello di ottenere per la geometria un insieme di assiomi completo (*vollständig*) e il più semplice possibile.

Il significato della dichiarazione d'intenti iniziale trova spiegazione nelle lettere a Frege.

> Non possiamo avere la definizione completa del concetto di punto prima che sia portata a termine la costruzione del sistema di assiomi. Ogni assioma contribuisce infatti alla definizione e

[11] E. V. Huntington, "A Complete Set of Axioms for the Theory of Absolute Continuous Magnitudes", *Trans. AMS*, vol. 3, 1902, pp. 265-79.

[12] O. Veblen, "A System of Axioms for Geometry", *Trans. AMS*, vol. 5, 1904, pp. 343-84.

[13] F. Enriques, *Per la storia della logica*, cit., pag. 198.

> ogni nuovo assioma determina dunque un mutamento nel concetto. "Punto" è sempre qualcosa di diverso, rispettivamente nella geometria euclidea, non-euclidea, archimedea, non-archimedea.
>
> L'introduzione di un qualunque assioma, una volta che un concetto sia stato stabilito in modo *completamente definito e univoco*[14], è qualcosa di non permesso e illogico: un errore che viene compiuto molto spesso, soprattutto da parte dei fisici. Dal fatto che essi creano a ripetizione sempre nuovi assiomi, senza porli a confronto con i precedenti, derivano vere e proprie assurdità[15]. Appunto nel procedimento di creare un assioma, di richiamarsi alla sua verità (?), e di concludere che lo stesso si accorda con i concetti definiti, sta la perpetua fonte di errori e malintesi.

Hilbert usa per questa proprietà sia il termine *vollständig* sia il termine "chiuso". Egli cerca nelle *Grundlagen* "un sistema di assiomi così chiuso che nessun nuovo assioma possa esservi coerentemente aggiunto, o che nessuna affermazione possa essere ritenuta corretta a meno che non sia derivabile dagli stessi".

Anche a proposito dei numeri reali presentati in modo assiomatico, nella conferenza del 1901, Hilbert afferma che per la loro introduzione non c'è più bisogno di pensare a tutte le leggi di generazione dei numeri, come nel metodo genetico, ma solo a "un sistema di cose le cui relazioni sono date mediante quel sistema finito e chiuso di assiomi I-IV su cui valgono nuove asserzioni solo se possono essere derivate da quegli assiomi per mezzo di un numero finito di passi logici".

Definire in modo "completo e univoco", soprattutto "univoco", un concetto fa pensare ovviamente alla categoricità. Ma la proprietà a cui Hilbert fa riferimento è un'altra: una teoria completa nel senso di Hilbert non ammette nuovi assiomi perché, come egli osserva ("nessun nuovo assioma possa esservi coerentemente aggiunto, o che nessuna affermazione possa essere ritenuta corretta a meno che non sia derivabile dagli stessi"), ogni enunciato o è già derivabile o è contraddittorio con la teoria.

Se un enunciato φ non è derivabile dagli assiomi T, allora non deve poter essere aggiunto coerentemente a T, cioè $T \cup \{\varphi\}$ deve risultare contraddittorio, e questo significa che $T \vdash \neg\varphi$.

[14]Corsivo nostro.

[15]Hilbert si riferisce a Hertz e alla teoria dell'elettromagnetismo.

Quindi si richiede che la teoria abbia la seguente proprietà: che per ogni enunciato φ, o φ o la negazione $\neg\varphi$ siano un teorema (non tutti e due per escludere la contraddittorietà). Questa è la proprietà che in seguito è stata chiamata *completezza* della teoria (per chiarezza ora la chiameremo completezza deduttiva). Il termine più usato per definirla all'inizio - quando la si contemplava nella sua chiara formulazione logica - era tuttavia, significativamente, quello di *Entscheidungsdefinitsheit*, che esprime la possibilità di una decisione: ogni enunciato può essere *deciso* nella teoria, nel senso che può essere o dimostrato o refutato (dimostrata la sua negazione)[16].

Conseguenza di questa proprietà è che un enunciato φ vero in una interpretazione è vero in tutte, perché è dimostrabile: altrimenti sarebbe dimostrabile la sua negazione, e questa sarebbe vera in tutte le interpretazioni, quindi φ non potrebbe essere vero in nessuna.

Anche Fraenkel, quando nella citazione sopra riportata celebra i vantaggi della categoricità, ha in mente la proprietà della completezza deduttiva, che comporta che un enunciato vero in una interpretazione è vero in tutte; infatti per lui in tale caso una proposizione valida per un determinato significato è "deduttivamente derivabile dagli assiomi", quindi valida per ogni significato.

Le due nozioni di categoricità e completezza deduttiva sono all'inizio confuse, sotto il termine "completezza" e identificate, con un interesse che si volge principalmente alla completezza deduttiva.

La desiderabilità della completezza deduttiva era duplice. Da una parte significava che ogni problema che si potesse formulare, nel linguaggio della teoria, era passibile, con gli strumenti messi a disposizione dalla teoria, cioè le dimostrazioni, di avere una soluzione, o positiva o negativa. Poincaré aveva chiaro anche che una conseguenza sarebbe stata una riduzione di ogni problema matematico a un problema logico, di ricerca di dimostrazioni formali.

L'altra aspettativa era appunto che siccome in questo caso tutte le interpretazioni avrebbero soddisfatto gli stessi enunciati e *per questo* esse sarebbero state indistinguibili a tutti gli effetti, quindi isomorfe, come si desiderava.

Invece risultò che le cose non stavano così. Per una teoria deduttivamente completa tutti i modelli sono in effetti indistinguibili, ma limitatamente alle proprietà esprimibili con formule del linguaggio della teoria, o "nell'ordine d'idee di cui si tratta" come dice Enriques; non è detto che siano isomor-

[16]Un enunciato che non è né dimostrato né refutato in una teoria si dice *indeciso* nella teoria, o *indecidibile* nella teoria.

fi[17]. Potrebbero esserci diversità non esprimibili nel linguaggio della teoria. Con linguaggi diversi magari si esprimono e si verificano proprietà non trasportabili da una interpretazione a un'altra.

Qualche dubbio era stato espresso da Huntington - sul rapporto tra categoricità e completezza deduttiva - in considerazione della nostra scarsa conoscenza della logica: "Nel caso di un insieme di assiomi categorico, si sarebbe tentati di asserire il teorema che se una proposizione può essere formulata in termini dei concetti fondamentali, allora o essa è deducibile dai postulati, oppure la sua negazione lo è: bisogna ammettere tuttavia che la nostra padronanza dei processi di deduzione logica non è ancora, e forse non sarà mai, sufficientemente completa" da chiarire la eventuale equivalenza[18].

Il teorema che tenta Huntington è vero, la categoricità implica la completezza deduttiva, è il contrario che non lo è.

Già nel 1901 Hilbert non sembrava più convinto della identità delle due nozioni: secondo una testimonianza di Edmund Husserl, in una conferenza in cui aveva discusso l'argomento Hilbert aveva osservato che il fattore cruciale per dirimere questioni di derivabilità era la logica con la quale si dimostrano le conseguenze degli assiomi. Scrive Husserl riassumendo: "quando supponiamo che una proposizione sia decisa sulla base degli assiomi di un dominio, cosa possiamo usare oltre a questi assiomi? *Alles Logische. Was ist das?* Tutte le proposizioni che siano libere dalle particolarità di un dominio di conoscenza, ciò che è indipendente da tutti gli assiomi particolari, da tutta la materia della conoscenza". Ma allora si ha un ventaglio di possibilità: "il dominio della logica algoritmica, il dominio del numero, il dominio della combinatoria, quello della teoria generale degli ordinali. E infine la teoria più generale degli insiemi non è essa stessa logica pura?" La logica combinatoria basta per derivare qualunque *Schnittpunktsatz* dal teorema di Pascal (Hilbert aveva notato che non serve la continuità), la logica del numero interviene quando si usa il postulato di Archimede e per usare il *Vollständigkeitsaxiom* (vedi oltre) occorre la logica degli insiemi, la *allgemeinste Mannigfaltigkeitslehre*.

Se è vero che non si sapeva ancora abbastanza di logica, ancora meno tuttavia si era in grado di dominare le manipolazioni di strutture con le loro

---

[17]Si diranno "elementarmente equivalenti", una nozione che non è molto familiare ai matematici, accanto a quella di isomorfismo e di omomorfismo, perché richiede un riferimento esplicito al linguaggio della teoria.

[18]E. V. Huntington, "A Set of Postulates for Ordinary Complex Algebra", *Trans. AMS*, vol. 6, 1905, pp. 209-29.

relazioni e costruzioni[19], e lo slittamento dalla nozione di categoricità, che era quella desiderata, alla completezza deduttiva, era forse anche dovuto alla possibilità di spostarsi sul terreno apparentemente più familiare dei teoremi e delle deduzioni.

La completezza a cui pensava Hibert per la geometria rispondeva comunque ad esigenze multiple e non ancora chiaramente articolate.

La completezza deduttiva della geometria euclidea in particolare si poneva come problema in relazione al rapporto tra intuizione spaziale ed assiomi della geometria piana, accennato in precedenza, e poteva essere un obiettivo perseguibile per se stesso.

Un sistema di assiomi deduttivamente completo per la geometria piana escluderebbe la necessità di ricorrere ad altri principi, ad altre conoscenze, ad esempio quelle relative all'immersione nello spazio, per dimostrare i teoremi del piano. Infatti se una teoria T è deduttivamente completa, allora se un enunciato si dimostra (magari più facilmente) in un'estensione non contraddittoria T_1, esso si dimostra anche (magari in modo meno agevole) nella teoria T; infatti altrimenti, per la completezza, nella teoria si dimostrerebbe la negazione,

$$\begin{array}{lll} T_1 & \vdash & \varphi \\ \bigcup & & \\ T & \vdash & \neg\varphi \qquad (\text{se } T \nvdash \varphi) \end{array}$$

e l'estensione T_1, che contiene tutti i teoremi di T, sarebbe contraddittoria.

Se si avesse una teoria deduttivamente completa, oltre ai vantaggi della completezza si avrebbe anche un potenziamento dei metodi dimostrativi per la teoria, potendosi ammettere anche quelli di ogni estensione non contraddittoria.

Questo è un vantaggio tecnico, ma non consiste solo in un ausilio euristico: grazie alla sua validità ad esempio, l'uso dei numeri in geometria non inficia l'autonomia della teoria stessa.

Nella sua ricerca di un sistema completo nel senso detto per la geometria, Hilbert formula tuttavia (a sorpresa?) anche un un assioma specifico di completezza. Il motivo può essere che non riteneva sufficiente, come per Poincaré, la certezza morale di essere riuscito ad assicurarla con gli altri assiomi; oppure incominciava a pensare che la completezza deduttiva poteva

[19]Nozioni come quella di omomorfismo ad esempio verranno più tardi.

dipendere dalla logica usata; oppure aveva presente le successive estensioni dei sistemi numerici e voleva imporre un esaurimento delle possibilità espansive; lui stesso probabilmente stava cercando di chiarirsi le idee con tentativi in diverse direzioni; un motivo plausibile è che siccome con il suo sistema chiuso di assiomi voleva assicurare la indipendenza della geometria dai numeri, Hilbert abbia pensato che fosse necessario un assioma che garantisse le stesse possibilità che l'uso dei numeri reali nei ragionamenti geometrici comportava. Fatto si è che Hilbert formula un *Vollständigkeitsaxiom* per la geometria che sembra essere una richiesta esplicita di categoricità: l'assioma afferma l'impossibilità di estendere l'universo dei punti, rette e piani e ancora mantenere la validità di tutti gli altri assiomi.

Per i reali allo stesso modo si postula l'impossibilità di un'estensione che conservi tutti gli altri assiomi: addio nuvolette di infinitesimi.

Questo assioma interferisce con quello di continuità, anzi la continuità è sostituita da Archimede e dalla *Vollständigkeit*.

L'assioma di *Vollständigkeit* è un assioma metateorico, e Hilbert si renderà presto conto, come risulta dal dialogo con Husserl, che esso richiede un linguaggio diverso.

Per quel che riguarda la terminologia attuale, si ricordi che ora la completezza del sistema dei numeri reali, quando si dice che formano un campo ordinato completo, non è espressa dal *Vollständigkeitsaxiom* hilbertiano ma dalla proprietà che ogni insieme limitato superiormente ha un estremo superiore, quella che una volta si chiamava continuità. Questa proprietà non ha nulla a che vedere con la completezza deduttiva di una teoria.

Quando si parla dei numeri reali, oggi con "completezza" s'intende l'assioma di completezza ora ricordato per i campi completi. Quando si parla di una teoria, con "completezza" s'intende la completezza deduttiva, l'*Entscheidungsdefinitsheit*. L'incompletezza di una teoria è la negazione di quest'ultima nozione, l'esistenza di un enunciato indecidibile nella teoria.

Per complicare ulteriormente il quadro terminologico, si ricordi che ancora diversa è la completezza della logica (teorema di Gödel del 1929), che afferma che un sistema di assiomi e regole è adeguato a dedurre tutte le formule logicamente valide.

Il programma di Hilbert, perseguito da lui e i suoi allievi[20] negli anni Venti, poneva l'obiettivo di dimostrare la non contraddittorietà e la completezza (*Entscheidungsdefinitisheit*) dell'aritmetica.

[20] Paul Bernays, Wilhelm Ackermann, John von Neumann.

4 La nozione di "formale"

Può sorprendere che un logico importante come Frege, uno dei padri della logica contemporanea, sia stato cieco di fronte allo sviluppo delle teorie assiomatizzate e così sordo alle ragioni di Hilbert. Nonostante abbia inventato l'ideografia, la scrittura logica interamente simbolica, Frege non era un logico formale; apparteneva piuttosto alla tradizione idealista, che affidava alla logica, o alla ragione, una capacità costitutiva, di portare cioè in essere, con una realtà oggettiva, i concetti che definiva.

Si ricordi che Pasch diceva in sostanza che se si sbaglia è per aver usato il significato, "tanto è vero che è proprio quando diventa necessario [pensare al significato] che si manifesta il carattere lacunoso della deduzione e (quando non si può sopprimere le lacune modificando il ragionamento) l'insufficienza delle proposizioni invocate come strumento di prova".

Perché con il formale non si sbaglia, le manipolazioni sintattiche sono meccaniche.

Il formale è stato scoperto dai greci, Aristotele e Stoici, ma non applicato alla matematica. La matematica antica, giustamente, non era formale, come non lo era quella dell'Ottocento, e come non lo è quella odierna, anche se ora si sa che in linea di principio lo è e lo deve essere.

La scoperta di Aristotele poi è stata svalutata, agli inizi dell'età moderna (gli Stoici erano stati dimenticati) perché i sillogismi erano un formalismo troppo povero per trattare gli argomenti matematici.

Se la logica s'identifica con la teoria dei sillogismi, non si può dar torto al rilievo sulla marginalità o mancanza totale di un suo ruolo nel ragionamento matematico. Ma l'irrilevanza dipende solo dalla povertà del linguaggio. Anche se si prende in considerazione il ragionamento proposizionale, non c'è molto della teoria logica che torni utile, salvo qualche tautologia che è alla base delle forme usuali di inferenza: il *modus ponens*, la riduzione all'assurdo, la contrapposizione, la distinzione di casi. Bisogna attendere la logica piena delle relazioni nel secondo Ottocento per realizzare l'intreccio di logica e matematica. Tuttavia dei quantificatori forse solo l'equivalenza tra $\neg\forall$ e $\exists\neg$ verrebbe citata come consapevole legge logica da parte di un matematico.

Quello che è rilevante della logica, anche di quella tradizionale, per la matematica, non sono le singole tautologie o le regole di inferenza, ma è la concezione stessa della logica, o dell'inferenza logica.

Per illustrare in cosa consista la logica formale, bisogna entrare nei dettagli di una specie di lezione introduttiva, perché si tratta di nozioni del tutto ignote ai più.

Nulla val meglio di un esempio. Dato il sillogismo

$$
\text{(I)}\qquad
\begin{array}{l}
\text{Nessun triangolo rettangolo è equilatero} \\
\text{Qualche triangolo isoscele è equilatero} \\
\hline
\text{Qualche triangolo rettangolo non è isoscele,}
\end{array}
$$

dove l'affermazione sotto la linea si chiama conclusione, e le due affermazioni sopra la linea premesse del sillogismo, si chiede se l'inferenza è corretta, o valida. Dopo una pausa di riflessione, raccogliamo le risposte, che sono frequentemente sbagliate, cioè sono risposte positive. Il motivo non è che non si sa ragionare, ma che non si capisce la domanda.

La domanda non è se la conclusione sia vera, ma se essa sia *conseguenza* delle due premesse. La questione è cosa s'intenda con "conseguenza". Si tratta di necessità metafisica, o di causa, o di costrizione mentale da parte di un demone, più forte di noi, che ci forza l'assenso, una volta accettate le premesse?

Le definizioni che si trovano nelle introduzioni dei manuali, anche contemporanei, sono tutte inficiate dalla necessità di essere brevi e dalla volontà di rendere facile una nozione che non è né semplice né intuitiva, anzi è una delle grandi scoperte del pensiero umano. Si trova detto che la relazione di conseguenza vale quando "se uno accetta le premesse, deve accettare la conclusione" - si sposta così il problema al significato di "deve"; si dice anche, indifferentemente, nonostante la disparità di piani di discorso (soggettivo l'uno, oggettivo l'altro), che la conseguenza sussiste quando "se le premesse sono vere, la conclusione è vera". La prima definizione è piuttosto da argomentazione giuridica, la seconda rinvia al problema equivalente di capire cosa vuol dire "se ... allora".

Il sillogismo

$$
\begin{array}{l}
\text{Nessun triangolo rettangolo è equilatero} \\
\text{Qualche triangolo scaleno è equilatero} \\
\hline
\text{Qualche triangolo rettangolo non è scaleno,}
\end{array}
$$

con una premessa falsa confonderebbe ulteriormente le idee, come ogni condizionale con antecedente falso.

Si dice anche, invece di "se ... allora": "*ogni volta* che le premesse sono vere, anche la conclusione è vera". Tuttavia non è chiaro cosa significhi "ogni volta" se le proposizioni non dipendono da un parametro: i triangoli rettangoli non sono equilateri, e non lo sono una volta per tutte, non qualche volta sì e qualche volta no.

Nel caso del sillogismo (I) una persona che capisca i termini del discorso e sappia di geometria accetta le premesse e accetta anche la conclusione. Non sussiste tuttavia una relazione di conseguenza. Per spiegarlo, la via più usuale e convincente è quella di osservare:

valido? certo che no, sarebbe come dire che

$$\text{(II)} \qquad \frac{\begin{array}{l}\text{Nessun cane è ruminante}\\ \text{Qualche quadrupede è ruminante}\end{array}}{\text{Qualche cane non è quadrupede}}$$

è un sillogismo valido.

Si tratta infatti dello stesso sillogismo - nel senso che se pure i predicati che vi compaiono sono diversi, le parole sono distribuite nello stesso modo, stabilendo tra i predicati corrispondenti le stesse relazioni - e in questo caso nessuno avrebbe alcun dubbio sulla non correttezza dell'inferenza.

Se si capisce e si accetta questa risposta, con il suggerimento implicito che contiene, vuol dire che si capisce che in verità il sillogismo di cui si parla non è né il primo proposto né quello portato a controesempio, ma

$$\text{(III)} \qquad \frac{\begin{array}{l}\text{Nessun } S \text{ è } M\\ \text{Qualche } P \text{ è } M\end{array}}{\text{Qualche } S \text{ non è } P}$$

cioè il sillogismo non riguarda né triangoli né cani, né altro, non è costituito da frasi che parlino di qualcosa. Il sillogismo è formato con schemi[1] di frasi, dove le lettere S, ... che vi compaiono non sono lettere dell'alfabeto ma variabili. La domanda sulla correttezza concerne la *forma* delle proposizioni.

Gli argomenti in linguaggio naturale, come quelli proposti in (I) e (II), sono solo esempi, o *interpretazioni* particolari, del sillogismo - delle sue variabili - con diversi insiemi di predicati.

Un argomento in linguaggio naturale[2] si dice corretto se è un caso particolare di un sillogismo formale valido.

L'interpretazione con i cani e i ruminanti mostra che il sillogismo (III) non è valido (e che (I) non è corretto). Se si accetta anche questa affermazione,

[1]La parola "schema" compare nella descrizione di Hilbert a Frege del metodo assiomatico.

[2]O in un formalismo interpretato, ad esempio una dimostrazione aritmetica.

significa che si capisce pure il secondo suggerimento implicito nella risposta, vale a dire la definizione di conseguenza logica.

Se un sillogiamo non è valido quando esiste un'interpretazione per cui le premesse sono vere e la conclusione è falsa (come è il caso di (II)), allora un sillogismo è valido quando ogni interpretazione che renda vere le premesse rende vera la conclusione.

Solo in riferimento alla forma con le variabili ha senso parlare di "ogni volta"; s'intende: "ogni volta che se ne dà una diversa interpretazione". Solo così si può dare quindi un senso alla definizione di conseguenza logica che si è imposta ed è stata accettata e viene insegnata.

Questa nozione di conseguenza logica ha una formulazione precisa ma appare assai difficile da verificare, per la necessità di prendere in considerazione la totalità infinita e neanche ben delimitata delle possibili interpretazioni.

Tuttavia non esiste altra definizione di conseguenza *logica* se non quella formale.

Ogni inferenza *logica* è in verità una metainferenza relativa a tutte le inferenze di una forma determinata che si possono incontrare nei discorsi dotati di significato.

Il matematico sa che quando un problema difficile ha una rappresentazione formale c'è speranza di poterlo dominare, lavorando sulla rappresentazione formale.

La teoria dei sillogismi formali è dovuta ad Aristotele. Anche gli Stoici avevano una concezione formale nel loro trattamento della logica proposizionale; essi esprimevano al esempio il *modus ponens* come "se il primo allora il secondo, ma il primo, dunque il secondo", usando "primo" e "secondo" per quello che poi si è scritto p e q, cioè variabili. Si possono usare parole senza senso, come "se pirlo allora fanto, ma pirlo, dunque fanto", o "se punto allora retta, ma punto, dunque retta".

La scoperta di Aristotele e degli Stoici è che esistono costrutti linguistici, che di solito coinvolgono le parole "se", "e", "non" (e per gli Stoici anche altre, come "o"), a cui le persone si affidano per accettare informazioni senza verificarle direttamente, subordinatamente all'accettazione di altre; tali costrutti sono frequenti e ricorrenti in molti contesti diversi. Non si sa - non si sa ancora oggi - come siano radicati nella mente, ma sono una componente importante del linguaggio evoluto; ed è possibile rendere coscienti questi schemi e farli oggetto di una considerazione metalinguistica che ne descriva la funzione generale per mezzo della considerazione della loro forma, espressa con un simbolismo in cui sono essenziali le variabili ($S, M, \ldots p, q \ldots$, primo, secondo, ...).

Solo la povertà del linguaggio a cui sono confinati i sillogismi, che discuteremo in seguito, li rende inadeguati a rappresentare i ragionamenti matematici. Ma l'idea della relazione di conseguenza è quella giusta e il legame con la matematica è stretto, se non addirittura di identità.

Le due premesse del sillogismo (III) si possono considerare assiomi di una mini teoria. La conclusione "Qualche S non è P" non è un teorema della teoria i cui assiomi sono "Nessun S è M" e "Qualche P è M". Abbiamo visto che non è conseguenza logica dei due assiomi; lo è invece

"Qualche P non è S".

La geometria propone assiomi con parole come "punto" e "retta", ma non parla di punti e rette come non parla di coppie di numeri e di equazioni lineari; allo stesso modo i sillogismi (I) e (II) non si riferiscono a triangoli e ad animali, almeno per quel che riguarda la domanda sulla validità, come è il caso con i teoremi. Le parole "punto" e "retta" potrebbero essere sostituite da X e Y, e il primo assioma di Euclide potrebbe diventare, in una formulazione più vicina al nostro modo usuale di esprimerci:

se a e b sono X esiste r che è Y e a e b sono gli estremi di r.

Anche questa semplice frase sarebbe già più complicata di quelle che possono comparire in un sillogismo. Non ci sono solo i predicati X e Y ma una relazione ternaria "a e b sono gli estremi di r".

4.1 Un modello matematico

Chi non fosse interessato alla logica può vedere la trattazione di questa lezione come un esempio di come si svolge un'indagine di matematica applicata. Queste iniziano con

1. Un problema Un problema reale, della vita concreta, che all'inizio è informe, non ben definito, soprattutto non matematico. Nel nostro caso il problema è quello di dare ragione della nozione di *conseguenza* e del suo uso, che si manifesta in molte dispute e argomentazioni, fin dall'origine della filosofia. Vi compare in modo non univoco, con una varietà di accezioni.

L'analisi matematica inizia con la costruzione di

2. Un modello La proposta di definizioni precise per fissare il problema, come si è fatto sopra con la definizione di conseguenza logica. Non sempre si riesce a catturare subito nelle definizioni tutte le possibili manifestazioni, diramazioni e appendici del problema. Bisogna fare delle semplificazioni, inizialmente anche drastiche, come vedremo in **4.1**.

Quindi si usa il modello per risolvere problemi ispirati dal campo concreto da cui si è partiti, e per la risoluzione di problemi occorrono

3. Le tecniche Formule, algoritmi. Il modello nel nostro caso si usa per riconoscere nella pratica l'uso corretto della relazione di conseguenza, quando un argomento particolare si conforma alle definizioni date, quindi si deve stabilire quali sono i sillogismi validi. Per questo cercheremo una tecnica in **4.2**, cercando possibilmente una tecnica effettiva, che sia (come *optimum*) un metodo di decisione meccanico, e vedremo i vantaggi di una tecnica matematica rispetto a una trattazione informale dei problemi.

Se le prime tecniche non sono completamente soddisfacenti, se ne cercano altre più sofisticate.

4. Altri modelli ed estensioni Nuove tecniche richiedono di solito altri modelli o, detto in altro modo, altre rappresentazioni dei dati del problema, che di solito ricoprono un ambito maggiore di fenomeni. Più articolata e ricca la rappresentazione, più spazio di manovra, e metodi più duttili. Nel nostro caso si passa dai sillogismi alle inferenze espresse nella logica del primo ordine, con connettivi e quantificatori, in **4.4**.

4.2 Sillogismi

Le frasi che intervengono nei sillogismi, studiati da Aristotele, dalla Scolastica e da tutta la logica fino a metà Ottocento, contengono solo predicati monadici, a un posto, e inoltre sono di una forma particolare.

Le frasi prese in considerazione e combinate tra loro affermano sempre una delle seguenti circostanze:

tutti quelli che hanno una proprietà P hanno anche la proprietà Q,

nessuno di quelli che hanno una proprietà P hanno la proprietà Q,

qualcuno che ha la proprietà P ha anche Q,

qualcuno con la proprietà P non ha Q.

Frasi di questa forma si chiamano, nella tradizione logica, proposizioni categoriche[3].

A prima vista si direbbe che il loro studio costituisca un'analisi dei quantificatori "qualcuno", "tutti", "nessuno", ma è forse più corretto dire che mette in evidenza e sfrutta alcune relazioni insiemistiche dell'inclusione e della intersezione, e precisamente

$$\begin{array}{lll} A \cap B = A, & \text{ovvero} & A \subseteq B \\ A \cap B \neq A, & \text{ovvero} & A \not\subseteq B \\ A \cap B = \emptyset, & \text{ovvero} & A \text{ e } B \text{ sono disgiunti} \\ A \cap B \neq \emptyset, & \text{ovvero} & A \text{ e } B \text{ non sono disgiunti.} \end{array}$$

Le proposizioni categoriche venivano scritte e lette nel seguente modo, dove a fianco mettiamo la versione insiemistica e quella logica moderna, scrivendo $P(x)$ per "x ha la proprietà P":

A	Tutti i P sono Q	$P \subseteq Q$	$\forall x(P(x) \to Q(x))$
E	Nessun P è Q	$P \cap Q = \emptyset$	$\forall x(P(x) \to \neg Q(x))$
I	Qualche P è Q	$P \cap Q \neq \emptyset$	$\exists x(P(x) \land Q(x))$
O	Qualche P non è Q	$P \not\subseteq Q$	$\exists x(P(x) \land \neg Q(x))$

Le lettere a sinistra sono la sigla con cui venivano indicate le corrispondenti proposizioni: A ed I sono proposizioni affermative, rispettivamente universale ed esistenziale; E e O sono negative, rispettivamente universale ed esistenziale.

Le parole "tutti", "nessuno", "qualche" diventavano superflue, sostituite dalla sigla prefissa; scrivendo "$A:\ P\ Q$" si intendeva la frase "tutti i P sono Q", con "$I:\ P\ Q$" la frase "qualche P è Q", e così via.

Siccome c'era sempre solo una variabile individuale, non si è sentita l'esigenza di indicarla. Il verbo era sempre il verbo "essere".

I sillogismi categorici[4] sono inferenze in cui occorrono due premesse e una conclusione, tutte e tre proposizioni categoriche, costruite con tre predicati, di cui uno inevitabilmente occorre due volte nelle premesse, per stabilire un legame, e si possono presentare ad esempio nel seguente modo:

[3] Si ammetteva un'estensione a proposizioni contenenti termini singolari come "Socrate" introducendo un predicato S per "socraticità", con un unico elemento, e traducendo "Socrate è mortale" con "Tutti quelli che hanno la proprietà S sono mortali".

[4] D'ora in avanti diremo soltanto "sillogismi".

$$\begin{array}{rcc} I: & P & M \\ \underline{E:} & \underline{S} & \underline{M} \\ O: & S & P. \end{array}$$

Il predicato S si chiama di solito *soggetto*, il predicato P *predicato* e il predicato M *termine medio*.

Un sillogismo è valido se la conclusione è conseguenza logica delle premesse.

Tutti i possibili sillogismi si caratterizzano per due aspetti; il primo è il tipo di proposizioni categoriche che nell'ordine lo costituiscono, si chiama *modo*, ed è rappresentato da tre lettere; ad esempio il modo del sillogismo di sopra è $I\ E\ O$.

Il secondo parametro è la disposizione dei tre predicati nelle tre proposizioni, e si chiama *figura*. Se si standardizza le presentazione, ad esempio chiedendo che la conclusione contenga sempre la coppia $S\ P$ nell'ordine - da cui il nome di questi due predicati - osservando che il termine medio deve sempre comparire in entrambe le premesse, altrimenti si vede facilmente che il sillogismo non è valido[5] - e che l'ordine delle premesse non è rilevante (per la terminologia, si chiama premessa *maggiore* quella che contiene S, e *minore* l'altra) si ottengono in tutto quattro figure:

$$\begin{array}{cccc} S\,M & S\,M & M\,S & M\,S \\ P\,M & M\,P & M\,P & P\,M. \end{array}$$

I sillogismi possibili sono 256 (esercizio), ma quelli validi solo[6] 15.

Come si fa a provare che un sillogismo è valido? È più facile mostrare che non lo è, come si è visto con (I), perché basta trovare un controesempio, vale a dire un'interpretazione in cui le premesse sono vere e la conclusione falsa.

Se un sillogismo è presentato in linguaggio naturale, un eventuale controesempio si può dare anche in linguaggio naturale, con altri predicati dotati di significato. Lo abbiamo visto con (I), dove si sarebbe anche potuto usare come controesempio

[5] A meno che la conclusione non coincida con una premessa.

[6] Una lunga controversia ha riguardato il fatto di considerare o no predicati vuoti, decisione che influenza la validità o meno di certi sillogismi. Nell'antichità si preferiva evitarli; nella trattazione moderna prevale l'interpretazione cosiddetta booleana, che ammette predicati vuoti. Sono 15 i sillogismi validi nell'interpretazione booleana.

$$\begin{array}{l}\text{Nessun cerchio quadrato è equilatero}\\ \underline{\text{Qualche triangolo è equilatero}}\\ \text{Quindi qualche cerchio quadrato non è un triangolo}\end{array}$$

conclusione che implica che esiste un cerchio quadrato.

Lo stessa ricerca di un controesempio in linguaggio naturale si può fare per un sillogismo formale. Ma non è mai del tutto chiaro che si sia esibito un controesempio, se gli esempi sono fatti in linguaggio naturale. Nessun predicato definito nel linguaggio naturale ha i confini esattamente delimitati senza ambiguità. Per questo i controesempi è meglio che siano insiemi astratti, dove gli elementi e i non elementi sono individuati in maniera precisa e indiscutibile.

L'interpretazione allora non è costituita da una terna di predicati dotati di senso, ma da insiemi astratti. Gli insiemi sono le estensioni di eventuali predicati (più uno, di solito indicato con U, che è l'universo di discorso).

Il linguaggio insiemistico per rappresentare interpretazioni funge da interfaccia tra quello simbolico e quello naturale e ha il vantaggio di esprimere in uno schema molte possibili interpretazioni concrete. Si tratta ancora di uno schema, come il sillogismo stesso, ma è intuitivamente più allettante, più vicino alle interpretazioni concrete, di cui costituisce una chiara indicazione per la definizione: qualunque scelta di predicati con quelle estensioni va bene.

Un secondo vantaggio dell'uso di insiemi è che lo studio delle interpretazioni, che costituisce la cosiddetta "semantica", si può impostare in modo matematico.

Entrambi i vantaggi si estendono a linguaggi e logiche più forti.

Nel caso in oggetto, un controesempio è fornito da $U = \{a, b\}$, con $S = \{a\}, P = \{a, b\}, M = \{b\}$.

Altri esempi:

Il sillogismo

$$\begin{array}{lcc} A: & S & M \\ \underline{I:} & \underline{M} & \underline{P} \\ I: & S & P \end{array}$$

è falsificato dalla seguente interpretazione: $U = \{a, b\}$, $S = \{a\}, M = \{a, b\}, P = \{b\}$.

Il sillogismo

$$
\begin{array}{ll}
E: & S \quad M \\
\underline{I:} & \underline{P \quad M} \\
O: & S \quad P
\end{array}
$$

non è valido se S è vuoto oppure se $S \subseteq P$.

Il sillogismo

$$
\begin{array}{ll}
A: & S \quad M \\
\underline{A:} & \underline{M \quad P} \\
I: & S \quad P
\end{array}
$$

non è valido in quanto se S è vuoto si ha $S \cap P = \emptyset$ e la conclusione è falsa, nonostante $S \subseteq M$ sia vero e $M \subseteq P$ si possa rendere vero con una scelta giusta di M e P[7].

Dimostrare la validità di un sillogismo è un altro paio di maniche: occorre fare un ragionamento che riesca a prendere in esame tutte le possibili interpretazioni. Ingegnosi argomenti si trovano nella tradizione logica. Ma un ragionamento si appoggia sempre alla logica, a qualche tipo di logica, per cui la certezza della validità del sillogismo si sposta alla questione della validità del ragionamento svolto - che non può essere del tutto banale perché le interpretazioni sono infinite.

Chi sia abituato a fare ragionamenti in forma moderna può verificare facilmente, in genere, che nei casi positivi la conclusione è deducibile dalle premesse, con passaggi del tutto naturali e di uso comune.

Ad esempio per il sillogismo

$$
\begin{array}{ll}
E: & P \quad M \\
\underline{I:} & \underline{S \quad M} \\
O: & S \quad P
\end{array}
$$

la dimostrazione è la seguente, dove in una colonna presentiamo l'argomento in linguaggio naturale ((i) e (ii) sono le premesse), nell'altra la sua versione

[7] Questo è un esempio di un sillogismo non valido nell'interpretazione booleana, mentre sarebbe valido se non fossero ammessi predicati vuoti.

formale, con la successione di formule logiche dei linguaggi moderni (su cui torneremo in seguito, e useremo queste deduzioni come esempi di applicazione delle regole logiche).

Linguaggio naturale	*Linguaggio logico predicativo*
(i) *Qualche S è un M*	1. $\exists x(S(x) \wedge M(x))$
Sia c un S che è anche un M	2. $S(c) \wedge M(c)$ *da* 1
Allora c è un S e	3. $S(c)$ *da* 2
c è un M	4. $M(c)$ *da* 2
(ii) *Nessun P è M*	5. $\forall x(P(x) \rightarrow \neg M(x))$
Se c è un P allora c non è un M, o *Se c fosse un P non sarebbe un M*	6. $P(c) \rightarrow \neg M(c)$ *da* 5
Se c è un M allora non è un P	*per contrapposizione da* 6 7. $M(c) \rightarrow \neg P(c)$
Ma siccome c è un M, *allora c non è un P*	*da questo e da* 4 8. $\neg P(c)$
Siccome c è un S, *c è un S che non è un P*	*da questo e da* 3 9. $S(c) \wedge \neg P(c)$
Qualche S non è un P	10. $\exists x(S(x) \wedge \neg P(x))$

Si può anche proporre, se si è più a proprio agio, un ragionamento insiemistico:

le premesse sono $P \subseteq {\sim} M$ e $S \cap M \neq \emptyset$; dalla prima segue $M \subseteq {\sim} P$, ma da

$$M \subseteq {\sim} P \text{ e } S \cap M \neq \emptyset$$

segue la conclusione $S \cap \sim P \neq \emptyset$.

Valido è anche il sillogismo

$$\begin{array}{lcc} A: & S & M \\ A: & M & P \\ \hline A: & S & P \end{array}$$

con la seguente dimostrazione:

Linguaggio naturale	*Linguaggio logico predicativo*
(*i*) *Tutti gli S sono M*	1. $\forall x(S(x) \to M(x))$
Uno che sia un S è anche un M	2. $S(x) \to M(x)$ *da* 1
(*ii*) *Tutti gli M sono P*	3. $\forall x(M(x) \to P(x))$
Uno che sia un M è anche un P	4. $M(x) \to P(x)$ *da* 3
Uno che sia un S è anche un P	*per transitività da* 2 *e* 4 5. $S(x) \to P(x)$
Tutti gli S sono P	6. $\forall x(S(x) \to P(x))$

In modo del tutto analogo, si dimostra la validità del sillogismo:

$$\begin{array}{lcc} A: & P & M \\ O: & S & M \\ \hline O: & S & P. \end{array}$$

Considerare un sillogismo alla volta non è soddisfacente; non finiremmo più[8]. Vorremmo una tecnica generale, uniforme, che non consista nel fare

[8] Aristotele è arrivato a prenderli in esame tutti sfrondando l'insieme con alcune considerazioni generali, ad esempio che nessuna conclusione generale può seguire da due premesse particolari, o di simmetria ($I : P \quad Q$ equivale a $I : Q \quad P$), e simili.

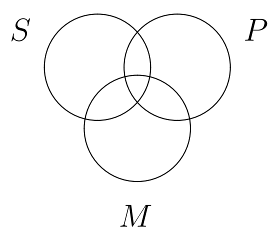
ogni volta una dimostrazione, o anche solo una riflessione.

Una tecnica matematica, quando la si trova, mostra tutti i vantaggi della matematica: rende quasi banale un problema che banale non è, riducendo a un solo caso - la dimostrazione che la tecnica funziona in modo corretto - quelli che costituirebbero una molteplicità intrattabile; dà più affidabilità dei singoli trattamenti perché è costituita di regole definite in modo netto, come manipolazione di simboli, e la manipolazione di simboli mostra la potenza del formale risolvendo con un numero finito di mosse la considerazione di una infinità di elementi.

Una tecnica risolutiva è tanto più attraente quanto più si riduce a una applicazione meccanica di regole meccaniche, che è l'aspirazione dietro ogni ricerca di un metodo risolutivo. Quello che si vorrebbe dunque per i sillogismi è un metodo di decisione, cioè un metodo effettivo uniforme che di fronte a un qualsiasi sillogismo permette di decidere se il sillogismo è valido o no.

4.2.1 Diagrammi di Venn

Esiste un metodo effettivo per determinare se un sillogismo sia valido o no. I tre predicati che intervengono in ogni sillogismo vengono rappresentati da tre insiemi, come in figura, senza escludere a priori nessuna possibilità per quanto riguarda le loro intersezioni (i loro complementi sono le regioni esterne ai cerchi).



Nel disegno sono presenti diverse aree: S, $S \cap P$, $S \cap M$, $S \cap P \cap M$, $S \setminus P$, $S \setminus M$, $S \setminus (P \cup M)$, ... Alcune possono essere vuote, altre no, a seconda degli insiemi, e di quanto su di essi stipulano le premesse.

Le premesse dei sillogismi si riferiscono solo ad alcune delle aree e sono equivalenti all'asserzione che certe intersezioni sono vuote o non vuote:

$$\begin{array}{rl}
A : S \quad M & S \cap (\sim M) = \emptyset \\
E : S \quad M & S \cap M = \emptyset \\
I : S \quad M & S \cap M \neq \emptyset \\
O : S \quad M & S \cap (\sim M) \neq \emptyset
\end{array}$$

e così le proposizioni categoriche che coinvolgono altre coppie di predicati.

Dato un sillogismo, prendiamo in esame le sue premesse (l'ordine non importa), senza guardare la conclusione. Se le aree corrispondenti ad una sua premessa devono essere vuote, perché la premessa sia vera, le tratteggiamo, se devono essere non vuote mettiamo una crocetta all'interno.

Oggettiviamo così in un disegno le informazioni date dalle premesse, dopo di che guardiamo soltanto se queste contengono l'affermazione della conclusione, che si riferisce sempre a un'area che deve risultare o tratteggiata o contenente una crocetta.

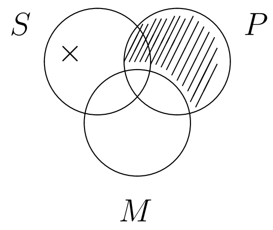
Ad esempio, per:

$$\begin{array}{lcc}
A: & P & M \\
O: & S & M \\
\hline
O: & S & P
\end{array}$$

la prima premessa ci fa tratteggiare $P \cap (\sim M)$, che deve essere vuota. La seconda premessa ci fa mettere una crocetta in $S \cap (\sim M)$.

Quest'area sarebbe composta di due parti, quella che interseca P e quella che non lo interseca; ma la prima è tratteggiata, in quanto $S \cap (\sim M) \cap P \subseteq P \cap (\sim M)$, già considerata; quindi la crocetta si mette in $S \cap (\sim M) \cap (\sim P)$[9].

[9] Se avessimo considerato per prima la seconda premessa, non avremmo avuto il vincolo che $S \cap (\sim M) \cap P = \emptyset$; allora la crocetta l'avremmo messa in mezzo, sulla linea di demarcazione tra $S \cap (\sim M) \cap P$ e $S \cap (\sim M) \cap (\sim P)$, come in un prossimo sillogismo (si veda oltre). Ma quando avessimo preso in considerazione la prima premessa, la crocetta sarebbe automaticamente caduta in $S \cap (\sim M) \cap (\sim P)$, o meglio l'avremmo noi spostata per chiarezza in quell'aerea.

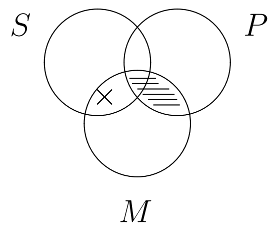


Ora si interpreta quello che dice la figura, come conseguenza delle due premesse, confrontandolo con quello che afferma la conclusione. La figura mostra una crocetta in $S \cap (\sim P)$, che è stata disegnata in quell'area, come conseguenza solo delle due premesse; è stata messa prima di guardare la conclusione. La conclusione afferma che deve esserci una crocetta in $S \cap (\sim P)$.

Un altro sillogismo valido è

$$\begin{array}{lcc} E: & P & M \\ \underline{I:} & \underline{S} & \underline{M} \\ O: & S & P \end{array}$$

La prima premessa ci fa tratteggiare l'area $P \cap M$ che deve essere vuota perché $P \subseteq \sim M$. La seconda premessa ci fa mettere una crocetta nell'area $S \cap M$, ma di fatto nell'area $S \cap M \cap (\sim P)$ perché $S \cap M \cap P = \emptyset$.

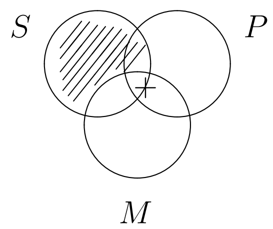


Come conseguenza si ha una crocetta in $S \cap (\sim P)$, che è quello che afferma la conclusione: $S \cap (\sim P) \neq \emptyset$.

Consideriamo ora un sillogismo non valido:

$$\begin{array}{lll} A: & S & M \\ \underline{I:} & \underline{M} & \underline{P} \\ I: & S & P \end{array}$$

La prima premessa fa tratteggiare $S \cap (\sim M)$ che è vuota essendo $S \subseteq M$. La seconda premessa ci fa mettere una crocetta in $M \cap P$, ma questa è divisa in due sottoaree $M \cap P \cap S$ e $M \cap P \cap (\sim S)$, e non sappiamo dove mettere la crocetta. La si mette sulla linea di divisione delle due sottoaree, proprio per indicare che non si sa in quale parte metterla.



Ma la conclusione dice che dovrebbe esserci una crocetta in $S \cap P$ e questa indicazione precisa non è conseguenza delle premesse; la crocetta è sul bordo, non è certamente nell'interno dell'area $S \cap P$, potrebbe essere in $(\sim S) \cap P$ e soddisfare comunque la seconda premessa. Le informazioni delle premesse non sono sufficienti. La conclusione non segue dalle premesse (dice di più).

Con i diagrammi di Venn non si ottiene un esplicito controesempio, quando il sillogismo non è valido, ma solo la certezza che c'è e indicazioni per definirlo, di solito abbastanza precise. Nel caso precedente ad esempio, si può facilmente pensare che se S fosse vuoto, o comunque disgiunto da P, la crocetta cadrebbe per forza in M, e non ci sarebbe nessuna crocetta in $S \cap P$.

Sarebbe preferibile un metodo che oltre a decidere se un sillogismo è valido fornisse anche un controesempio, quando il sillogismo non è dichiarato valido, e di nuovo possibilmente in modo meccanico. Metodi del genere esistono, ma non sono un'estensione dei diagrammi di Venn, richiedono altre considerazioni.

4.3 La logica formale moderna

La conclusione della riflessione moderna sul metodo assiomatico è stata che le teorie matematiche non parlano di niente, così come le proposizioni della logica formale, perché *le loro frasi non sono frasi ma schemi formali di frasi*; e se parlano di qualcosa, parlano di formule (come quando si dice che (III) non è valido); è un po' dura da mandare giù forse, la prima volta che lo si sente dire, ma la consolazione è che le teorie non parlano di niente perchè parlano di tutto, di tutto quello che può soddisfare gli assiomi.

Per rendersi conto in modo definitivo e accettare in modo cosciente tale natura del pensiero matematico, già riconosciuta nel suo ambito dalla logica formale, occorreva anche che la matematica potesse essere davvero formulata in modo formale. Questo fu reso possibile dall'arricchimento dei linguaggi logici realizzato dalla fine dell'Ottocento, da G. Frege e G. Peano, ben al di là del formalismo algebrico di Boole, ancora legato alla logica sillogistica.

I linguaggi logici moderni sono fondati su una diversa analisi linguistica rispetto alla sillogistica, è questa che li caratterizza, non l'uso massiccio o totale del simbolismo: mettono in evidenza individui e soprattutto relazioni tra essi, invece che solo predicati, e usano i quantificatori e tutti i connettivi; essi sono in grado ora di esprimere e rappresentare qualunque discorso, di quelli dichiarativi che si fanno in matematica, e non solo gli elementari legami tra predicati delle proposizioni categoriche.

4.3.1 Linguaggi predicativi

I linguaggi più usati e studiati, e le cui proprietà sono meglio conosciute, e più numerose, sono i cosiddetti linguaggi predicativi, o del primo ordine[10].

Le formule elementari dei linguaggi predicativi rappresentano le affermazioni che attribuiscono a uno o più soggetti la sussistenza di una proprietà o di una relazione.

Si usano simboli come $P, Q, \ldots$ per i predicati (o proprietà) e $R, S, \ldots$ per relazioni, e le *formule atomiche* sono tipicamente del tipo

$$P(t), R(t_1, t_2), S(t_1.t_2, t_3), \ldots$$

[10]Sono detti così perché hanno solo variabili individuali, e non di tipi superiori.

dove i t_i sono *termini*, che rappresentano individui, e sono in generale composti: sono o *variabili* (individuali)[11] $x, y, \ldots$, o *costanti* $c, d, \ldots$ o si ottengono applicando simboli di *funzione* a termini: $f(x), g(c, y), g(x, f(c)), \ldots$

Le formule si ottengono applicando ripetutamente connettivi $(\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \ldots)$ e quantificatori $(\forall, \exists)$[12] a partire da formule atomiche, ossia iterando l'applicazione delle seguenti clausole induttive:

1. se φ è una formula, anche $(\neg\varphi)$ è una formula,
2. se φ e ψ sono formule e $\bullet$ un connettivo binario, anche $(\varphi \bullet \psi)$ è una formula,
3. se φ è una formula e x una variabile, anche $(\forall x\varphi)$ e $(\exists x\varphi)$ sono formule.

I connettivi corrispondono alle funzioni di verità, che sono infinite, ma hanno basi finite: per comodità di scrittura si usano di solito quelli sopra indicati, ma teoricamente sono sufficienti le basi minimali $\{\neg, \wedge\}$ o $\{\neg, \vee\}$ o $\{\neg, \rightarrow\}$.

Lo studio della sintassi è lungo e complicato anche per una lingua naturale - si pensi ai plurali o alle desinenze dei verbi - nonostante le regole siano sempre incomplete, figuriamoci con una lingua dove non devono sussistere lacune o ambiguità. Il motivo di questa stringente richiesta può non essere ovvio, né univoco; i primi logici contemporanei, come Frege e Peano, vedevano nella precisione assoluta una garanzia di correttezza e certezza; altri volevano considerare la logica con gli stessi criteri di un sistema algebrico, nel suo *uso*; oggi si preferisce pensare che lo *studio* dei sistemi logici vada condotto con il rigore matematico.

Non ci soffermiamo sui molti dettagli sintattici necessari; ricordiamo ad esempio che le parentesi sono necessarie per permettere il riconoscimento delle formule nell'insieme di tutte le parole sull'alfabeto, e il loro *parsing*, o analisi sintattica (che individua il segno logico principale, le sottoformule e via ricorrendo); ma per evitare una loro proliferazione si introducono convenzioni sulla maggiore o minore forza attrattiva dei simboli logici, analoghe a quelle per le operazioni algebriche. Per le poche formule che dovremo considerare, o che abbiamo già considerato in **4.2.1**, la struttura sarà trasparente, grazie nel caso a parentesi ausiliarie[13]

[11] Ce ne è sempre una lista infinita; gli altri simboli, di costante, funzione e relazione, differiscono invece da linguaggio a linguaggio e possono anche non esserci.

[12] Anche questi simboli logici sono presenti in tutti i linguaggi.

[13] Si ricordi comunque che l'ordine di priorità è: $\neg, \forall, \exists$ al primo posto, quindi nell'ordine di forza decrescente $\wedge, \vee, \rightarrow, \leftrightarrow$.

Così il primo assioma di Euclide, discusso all'inizio di **4**, viene formalizzato, se con E si indica la relazione ternaria "x e y sono gli estremi di z", da

$$\forall x \forall y (X(x) \land X(y) \to \exists z (Y(z) \land E(x, y, z))).$$

Altri esempi, o esercizi di formalizzazione, si possono presentare con frasi del linguaggio naturale, o con frasi aritmetiche, dove il linguaggio non necessita di essere esattamente delimitato, e si possono usare le solite notazioni per le operazioni e le relazioni (in particolare con la loro posizione infissa invece che prefissa[14]). Per abbreviare la lunghezza delle espressioni si introducono simboli definiti di complessità incrementale.

x divide y

$$x \mid y \leftrightarrow \exists z (x \cdot z = y)$$

x è un numero primo

$$pr(x) \leftrightarrow x > 1 \land \forall y (y < x \to (y \mid x \to y = 1))$$

Esistono infiniti primi

$$\forall x \exists y (x < y \land pr(y))$$

Di due numeri consecutivi, uno è pari

$$\forall x \forall y (y = x + 1 \to \exists z (2z = x) \lor \exists z (2z = y))$$

oppure

$$\forall x \exists y (2y = x \lor 2y = x + 1)$$

Se un numero è pari, il suo successore è dispari

$$\forall x (\exists y (2y = x) \to \neg \exists y (2y = x + 1))$$

oppure

$$\forall x (\exists y (2y = x) \to \forall z (2z \neq x + 1))$$

[14] La notazione prefissa ha il vantaggio di rendere superflue le parentesi ai fini dell'analisi sintattica delle espressioni.

Chi ha un amico è felice[15]

$$\forall x(\exists y A(x,y) \to F(x))$$

oppure

$$\forall x \forall y(A(x,y) \to F(x))$$

Chi ama perdona

$$\forall x \forall y(A(x,y) \to P(x,y))$$

Maria ama il padre di Giovanni

$$A(m, f(g))$$

Maria ama solo il padre di Giovanni

$$A(m, f(g)) \wedge \forall x(A(m,x) \to x = f(g))$$

oppure

$$\forall x(A(m,x) \leftrightarrow x = f(g))$$

Maria ama un figlio di Giovanni

$$\exists x(F(x,g) \wedge A(m,x))$$

Maria ama i calciatori

$$\forall x(C(x) \to A(m,x))$$

[15]Per formalizzare frasi del linguaggio naturale, occorre decidere quali simboli utilizzare; in questo caso un simbolo di relazione $A(x,y)$ per "x e y sono amici" e un simbolo di predicato F per "essere felice". Si potrebbe anche decidere di usare un predicato $A(x)$ per "x ha un amico", ma si perderebbe qualcosa della struttura della frase; sarebbe impossibile aggiungere in segbuito qualche osservazione sugli eventuali amici di x.

Si noti come si realizzano diversamente gli articoli: il determinativo "il" con una descrizione, un termine; l'indeterminativo "un" e il plurale "i" con una variabile ma nel primo caso quantificata esistenzialmente e nel secondo universalmente.

La nozione più importante è quella di variabile libera e di variabile vincolata. Se non si sono quantificatori in una formula, tutte le variabili sono libere, in tutte le loro occorrenze. Se si premette un quantificatore a φ, in $\forall x\varphi$ o $\exists x\varphi$ tutte le occorrenze libere di x in φ diventano vincolate (dal quantificatore iniziale), mentre quelle di altre variabili restano quello che erano in φ. Se si combinano formule con un connettivo, lo stato delle variabili rimane quello che era.

Per indicare che una formula φ ha una variabile libera x si scrive anche come abbreviazione $\varphi(x)$[16]. $\varphi(t)$ indica il risultato della sostituzione del termine t a tutte le occorrenze libere di x in φ.

Tutti sanno dall'esperienza con la notazione funzionale cosa vuol dire scrivere $\varphi(t)$, ma a questo punto di un manuale di logica viene data di solito una definizione che appare complicata e del tutto superflua. Prima si definisce cosa si intende con "il risultato della sostituzione di t a tutte le occorrenze di x in un termine s", che si indica con $s[t/x]$ (e propriamente, per lasciare traccia della variabile che è stata sostituita, e non usare le parentesi tonde che fanno parte del linguaggio.). Per induzione sulla costruzione di t:

1. se s è x, $s[t/x]$ è t;
2. se s è una variabile diversa da x, $s[t/x]$ è s;
3. se s è una costante c, $s[t/x]$ è s;
4. se s è $f(t_1, \ldots, t_n)$, $s[t/x]$ è $f(t_1[t/x], \ldots, t_n[t/x])$.

Quindi si definisce $\varphi[t/x]$, risultato della sostituzione di t a tutte le occorrenze libere di x in φ, per induzione sulla costruzione di φ:

1. se φ è $R(t_1, \ldots, t_n)$, $\varphi[t/x]$ è $R(t_1[t/x], \ldots, t_n[t/x])$;
2. se φ è $\neg\psi$, $\varphi[t/x]$ è $\neg\psi[t/x]$;
3. se φ è $\psi \wedge \chi$, $\varphi[t/x]$ è $\psi[t/x] \wedge \chi[t/x]$;
4. se φ è $\exists x\psi$, $\varphi[t/x]$ è φ;

[16] Che non esclude che ne abbia anche altre.

5. se φ è $\exists y\psi$, y diversa da x, $\varphi[t/x]$ è $\exists y\psi[t/x]$.

Perché queste definizioni barocche e apparentemente ispirate a un atteggiamento formalista deteriore? Perché è importante mostrare come tutte le manipolazioni sintattiche che corrispondono ai comuni processi linguistici e inferenziali sono definibili in modo matematico - in questo e nella maggior parte dei casi in particolare in modo induttivo. Questo sarà importante per la dimostrazione dell'incompletezza.

Prima ancora, per mostrare che i comuni processi linguistici e inferenziali sono guidati dalla sintassi (*syntax driven*), anche quando magari si pensa che sia il (buon) senso che comanda.

Ad esempio si pensa che a una variabile libera x, che ha un significato universale, si possano sostituire termini t qualunque: da una formula $\exists y(x \neq y)$ si deduce $\exists y(t \neq y)$, dove t può essere un c, o $f(c)$, o $f(z)$ o altro, ma nessuno sostituirebbe y, perché $\exists y(y \neq y)$ ha tutt'altro senso, in particolare è ovviamente falso. Le restrizioni su t possono essere precisate in termini esclusivamente sintattici.

In una formula possono esserci sia occorrenze libere sia occorrenze vincolate di una variabile, come la x in

$$P(x) \wedge \forall x\varphi(x),$$

dove solo la prima occorrenza di x è libera. Se a questa formula si premette un quantificatore per x, ad esempio

$$\exists x(P(x) \wedge \forall x\varphi(x)),$$

solo la x in $P(x)$ risulta vincolata *dal* quantificatore iniziale:

$$\exists \underline{x}(P(\underline{x}) \wedge \forall x\varphi(x)).$$

In questo caso per evitare difficoltà di lettura all'occhio umano si esegue quella che si chiama una rinomina delle variabili, che produce formule equivalenti, e si scrive

$$\exists x(P(x) \wedge \forall y\varphi(y)).$$

Una formula che non ha occorrenze libere di nessuna variabile si chiama enunciato.

4.3.2 Semantica

Le interpretazioni di un linguaggio (determinato dai simboli predicativi, relazionali, funzionali e di costante che ha nell'alfabeto) sono presentate direttamente in termini di insiemi (come si è già visto per i sillogismi) e sono strutture $\mathcal{M}$ con universo M non vuoto e tali che in esse sono definiti

- un insieme $P^{\mathcal{M}} \subseteq M$ per ogni simbolo di predicato P,
- una relazione $R^{\mathcal{M}} \subseteq M^n$ per ogni simbolo di relazione n-aria R,
- una funzione $f^{\mathcal{M}} : M^n \to M$ per ogni simbolo di funzione n-aria f,
- un elemento $c^{\mathcal{M}}$ per ogni costante c.

A ogni termine t viene associata una funzione

$$t^{\mathcal{M}} : M^r \to M,$$

dove r è il numero delle variabili in t; se t non ha variabili, $t^{\mathcal{M}}$ è un elemento di M; se alle variabili $x_1, \ldots, x_r$ di t sono assegnati elementi[17] $a_1, \ldots, a_r$ di M, $t^{\mathcal{M}}(a_1, \ldots, a_r)$ è un elemento di M.

La relazione di *soddisfazione* è una relazione ternaria tra strutture, formule e assegnazioni di valori alle variabili.

Una formula atomica $R(t_1, \ldots, t_n)$, $1 \leq n$, è soddisfatta nell'interpretazione, una volta che si assegnino elementi $a_1, \ldots, a_r$ di M alle sue variabili - soddisfatta *dagli* elementi assegnati alle variabili - se gli elementi $t_1^{\mathcal{M}}(a_1, \ldots, a_r), \ldots, t_n^{\mathcal{M}}(a_1, \ldots, a_r) \in M$ stanno nella relazione (o predicato) $R^{\mathcal{M}}$.

Per sottolineare il fatto che la relazione di soddisfazione sussiste tra oggetti astratti ed è una relazione matematica, si introduce un simbolo speciale[18] $\models$ da usarsi in complessi del tipo $\mathcal{M} \models \varphi[\overline{a_i}]$ e da leggersi "in $\mathcal{M}$ φ è soddisfatta da $a_1, \ldots, a_r$"; la precedente definizione assume allora l'aspetto

$$\mathcal{M} \models R(t_1, \ldots, t_n)[\overline{a_i}] \quad \text{se e solo se} \quad \langle t_1^{\mathcal{M}}(\overline{a_i}), \ldots, t_n^{\mathcal{M}}(\overline{a_i})\rangle \in R^{\mathcal{M}}.$$

[17] Non necessariamente distinti.

[18] Non ne faremo molto uso nel seguito, anche se interviene, oltre che nella soddisfazione, in tutte le nozioni semantiche derivate, con estensioni del suo significato e qualche variazione notazionale.

Formule più complicate sono soddisfatte o meno a seconda del significato dei connettivi e dei quantificatori che vi intervengono, a partire dalle formule atomiche.

Non deve sorprendere che nella definizione della semantica si faccia appello al significato delle particelle logiche; la natura formale del ragionamento dipende dal fatto che i predicati, le relazioni e le funzioni non hanno un significato, sono *dummies* in attesa di essere scambiati con cose sostanziose[19]. Non vogliamo fondare il ragionamento umano, ma costruire uno strumento che agevoli, o costringa all'operazione di svuotamento di significato delle parole richiesto dal metodo assiomatico. Le particelle (in quanto distinte dalle parole) restano quelle solite.

Per le affermazioni atomiche, la corrispondenza tra

$$R(t_1, \ldots, t_n) \text{ è soddisfatta da } \overline{a_i} \qquad \text{e} \qquad \langle t_1^{\mathcal{M}}(\overline{a_i}), \ldots, t_n^{\mathcal{M}}(\overline{a_i})\rangle \in R^{\mathcal{M}}$$

esprime uno sdoppiamento tra frase e fatto che nel parlato è impedito dalla coincidenza tra il linguaggio nel quale si fanno affermazioni e il linguaggio nel quale si descrivono i fatti che rendono vere quelle affermazioni.

Se si pronuncia "Giovanni cammina" si esprime un fatto, ma il fatto che rende vera la frase "Giovanni cammina" è descritto dicendo ancora "Giovanni cammina".

Per separare i due fenomeni, quello linguistico e quello fisico, occorre introdurre un linguaggio che parli del linguaggio, e non dei fatti fisici, o non solo. Quando si spiega

"Giovanni cammina" è vero   se e solo se   Giovanni cammina

si sta usando a sinistra un linguaggio che parla del linguaggio - le frasi tra virgolette non sono frasi ma oggetti - a destra un linguaggio che parla di fatti. Per le lingue naturali i due linguaggi, quello di cui si parla e quello che parla dei fatti, coincidono, mentre nello studio logico è bene separarli, anzi necessario perché il linguaggio formale (quello nel quale c'è R) non parla di nulla.

In seguito si potranno di nuovo sovrapporre, ma con la sicurezza di tutta la sapienza acquisita.

Una formula è valida in una struttura se è soddisfatta da tutte le assegnazioni di valori alle sue variabili libere[20]. Una formula è soddisfacibile in una struttura se esiste almeno una assegnazione che la soddisfa.

Gli enunciati sono soddisfatti o no indipendentemente dalla assegnazione di valori alle variabili, che non compaiono libere, e si dicono rispettivamente

[19] L'immagine è dovuta al logico W. O. Quine.

[20] In tal caso si scrive $\mathcal{M} \models \varphi$.

veri o falsi nell'interpretazione. Un enunciato $\forall x\varphi(x)$ è vero in una struttura se e solo se $\varphi(x)$ è valida. Un enunciato $\exists x\varphi(x)$ è vero in una struttura se e solo se $\varphi(x)$ è soddisfacibile nella struttura.

Una formula con variabili libere non è né vera né falsa in una struttura, ma *definisce* l'insieme degli elementi (o delle coppie, ...) che la soddisfano.

$\forall x\exists y(x < y \wedge pr(y))$ è vero in $\mathbb{N}$ perché ogni n soddisfa $\exists y(x < y \wedge pr(y))$; qualunque sia n, tra tutti i numeri maggiori di n ce ne è uno che soddisfa $pr(y)$. Non è facile trovarlo, la verifica diretta delle affermazioni semantiche non è per nulla semplice.

Per quanto nel fare matematica sembri che si lavori sul piano semantico, in verità si utilizzano in continuazione conoscenze accumulate per abbreviare la verifica[21], o per renderla superflua. In questo caso si sa che basta cercarlo fino a $n! + 1$, ma si sa anche che è inutile cercarlo perché è dimostrato che esiste. La dimostrazione è più veloce e sicura che la ricerca in $\mathbb{N}$, e non vale solo per $\mathbb{N}$.

Una formula è logicamente valida se è valida in ogni interpretazione; un enunciato è logicamente vero se è vero in ogni interpretazione[22].

Se φ è vero (valida) in $\mathcal{M}$ si dice che $\mathcal{M}$ è modello di φ. Se $\mathcal{M}$ è modello di φ per ogni $\varphi \in T$ si dice che $\mathcal{M}$ è modello di T. Quindi si definisce

"φ è conseguenza logica di T" se ogni modello di T è modello di φ.

Per enunciati φ e ψ si ha che

"φ è conseguenza logica di ψ" se e solo se "$\psi \to \varphi$ è logicamente vero".

Conseguenza logica e validità logica sono legate in modo che le proprietà dell'una si trasportano all'altra.

Questa è la nozione di conseguenza semantica.

4.3.3 Derivazioni

Una nozione di conseguenza sintattica, o deduttiva, si ha introducendo la nozione di "derivazione da T", che è una successione finita di formule ciascuna o in T o ottenuta dalle precedenti per applicazione di una delle regole logiche. T è l'insieme degli assiomi: in ogni derivazione da T possono occorrere un numero finito di elementi di T che si chiamano propriamente assunzioni della derivazione.

[21] In un universo infinito, la verifica potrebbe essere un affare infinito.

[22] In tal caso si scrive $\models \varphi$.

Le regole deduttive codificano quello che si fa usualmente nelle dimostrazioni, aggiungendo solo precisazioni sintattiche che non stiamo a ricordare tutte perché le si applicano sempre naturalmente, o meglio con la sensibilità che ci dice che se non le rispettassimo si stravolgerebbe il senso delle frasi. Abbiamo già ricordato le cautele sulla sostituzione di termini.

Si formulano due regole per il quantificatore universale e due per il quantificatore esistenziale, rispettivamente una di introduzione e una di eliminazione.

L'eliminazione del quantificatore esistenziale è la regola di particolarizzazione

$$\frac{\forall x \varphi(x)}{\varphi(t)}$$

con restrizioni su t accennate sopra, e quella dell'introduzione del quantificatore esistenziale è l'indebolimento

$$\frac{\varphi(t)}{\exists x \varphi(x)}$$

L'introduzione dell'universale richiede la condizione che nella premessa della regola la x libera che si vuole generalizzare, o quantificare unversalmente, non sia libera nelle assunzioni delle derivazione:

$$\frac{\varphi(x)}{\forall x \varphi(x)}$$

Se ad esempio si inizia una derivazione con l'assunzione $x > 0$ e alla fine utilizzando gli assiomi dei campi ordinati completi si arriva a $\exists y(x = y^2)$, non si può applicare la generalizzazione e affermare $\forall x \exists y(x = y^2)$. La conclusione non vale per tutti gli x ma solo per quello che sono > 0 come affermato dalla premessa. Formalmente questa conclusione corretta si ottiene prima trasformando in modo legittimo (giustificato dal complesso del sistema di regole) la derivazione in una derivazione senza alcuna assunzione, oltre agli assiomi, di

$$x > 0 \to \exists y(x = y^2)$$

che vale per *tutti* gli x, e quindi introducendo il quantificatore universale

$$\forall x(x > 0 \to \exists y(x = y^2)).$$

L'eliminazione del quantificatore esistenziale è la più complessa: dato $\exists x\varphi(x)$ si introduce un nuovo simbolo c, costante o una variabile che non interferisca con altre parti precedenti della dimostrazione, quindi ancora non usata, e si scrive $\varphi(c)$ - e si dice "sia c un elemento tale che φ" - ma tale simbolo deve scomparire prima della fine della derivazione, per esempio con una applicazione dell'introduzione di $\exists$.

Le regole per i quantificatori non sono indipendenti, perché vale una relazione che definisce l'uno in termini dell'altro: per ogni φ la formula

$$\forall x\varphi \leftrightarrow \neg\exists x\neg\varphi$$

è logicamente valida, e derivabile dal complesso delle regole logiche. Formule logicamente valide si chiamano anche *leggi logiche*.

Quindi analogamente si danno regole di introduzione ed eliminazione per ciascun connettivo. Quella per l'introduzione del condizionale l'abbiamo vista sopra: se da $\{\varphi\} \cup T$ si deriva ψ, da T si derivi $\varphi \to \psi$[23]. Quella di eliminazione del condizionale è la classica regola del *modus ponens*, o di distacco:

$$\frac{\varphi \qquad \varphi \to \psi}{\psi.}$$

Ma il modo più veloce di completare il sistema logico è quello di dire che si possono usare come assunzioni tutte le tautologie, insieme con la regola *modus ponens*.

Ne vengono così alcune *regole derivate*: ad esempio siccome

$$(\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$$

è una tautologia, è giustificata la regola di contrapposizione

$$\frac{\varphi \to \psi}{\neg\psi \to \neg\varphi.}$$

[23]Questa regola appare un po' diversa dalla caratterizzazione generale di una regola sintattica, secondo la quale si guarda il segno logico della o delle formule già ottenute e a seconda di quale sia si sceglie la regola applicabile; nel caso dell'introduzione del condizionale occorre in un certo senso volere pervenire a $\varphi \to \psi$. Di conseguenza si decide di aggiungere una assunzione apposita. Tuttavia questo tipo di regole, che sono presenti nei cosiddetti sistemi di deduzione naturale, e non in altri, non modifica il discorso generale sul carattere effettivo delle derivazioni.

La tautologie sono particolari leggi logiche, cioè formule che sono logicamente valide, ottenute applicando i connettivi a formule che sono o atomiche o iniziano con un quantificatore, e tali che se si assegnano in tutti i modi possibili i valori 0 e 1 alle formule componenti e si calcola il valore secondo le funzioni di verità assegnate ai connettivi si ottiene sempre 1.

Siccome è decidibile in modo effettivo se una formula è una tautologia, si possono usare tutte, quando necessario, e la nozione di derivazione resta decidibile in modo effettivo, supposto che anche T sia decidibile. Si possono usare anche altre leggi logiche che non siano tautologie, se si riconosce che sono logicamente valide. Ma non è decidibile in generale se una formula è logicamente valida, come vedremo[24].

Queste regole si vedono in atto nelle derivazioni relative ai sillogismi che abbiamo considerato sopra. Come già anticipato, scriviamo $T \vdash \varphi$ per "φ è derivabile da T".

Quale vantaggio c'è a formalizzare le regole logiche? Uno è che oggettivando in regole le inferenze che di solito sono eseguite nel flusso del linguaggio è possibile riflettere sul modo come si ragiona. Le dimostrazioni formalizzate in derivazioni permettono un lavoro di sperimentazione e verifica, e di esercizio, che aiuta a familiarizzarsi anche con i ragionamenti meno intuitivi.

Ad esempio Saccheri chiamò *consequentia mirabilis* una forma di argomentazione che credeva di aver scoperto lui, mentre era stata usata, ma raramente (una volta in Euclide), e che consiste nello stabilire φ in base al fatto che φ segue dalla propria negazione $\neg\varphi$.

Si vede subito che

$$(\neg\varphi \to \varphi) \to \varphi$$

è una tautologia che giustifica l'inferenza; ma anche che la *consequentia mirabilis* può essere ricondotta alla più familiare dimostrazione per assurdo.

Infatti se si assume $\neg\varphi \to \varphi$ le si può associare la ovvia legge dell'identità $\neg\varphi \to \neg\varphi$, e da queste due si ottiene $\neg\varphi \to \varphi \wedge \neg\varphi$; siccome $\neg\varphi$ implica una contraddizione, si può affermare $\neg\neg\varphi$, cioè φ[25].

Ma l'interesse maggiore della formalizzazione sta nella possibilità di uno studio metateorico in grande, come vedremo nell'ultima lezione.

[24] Alcuni sistemi logici presentano un insieme finito di regole e un insieme decidibile, di solito finito, di leggi logiche, chiamate assiomi logici, che si suppone vengano sempre aggiunte a T; altri sistemi, quelli di deduzione naturale, hanno solo regole.

[25] Per simili esercizi di analisi delle dimostrazioni si veda G. Lolli, *QED Fenomenologia della dimostrazione*, Bollati Boringhieri, Torino, 2005, cap. 3.

5 Incompletezza e indecidibilità

Nel 1930 Gödel dà risposta negativa a entrambi i problemi della completezza e della dimostrabilità della non contraddittorietà dell'aritmetica.

Una volta capita l'importanza di questi risultati per la problematica che abbiamo descritto, ricostruendo il contesto storico, il loro interesse duraturo e attuale non si giustificherebbe, se non ci fosse una loro proiezione in avanti, con altre ragioni di apprezzamento.

Tali ragioni consistono non tanto nel contenuto dei loro enunciati quanto nella loro dimostrazione, nei concetti e nelle tecniche innovative ivi adottate, e che sono state poi ulteriormente sviluppate nella prosecuzione delle ricerche, sia da parte di Gödel stesso che da parte di altri logici.

In modo analogo ad esempio, la determinazione del valore di aree e volumi fatte da Eudosso e Archimede ci interessa meno del metodo di esaustione da loro applicato.

Per l'aritmetica, Gödel prova che non si può dimostrare la sua non contraddittorietà se non con strumenti più forti dell'aritmetica.

Cosa significa parlare della forza degli strumenti *matematici* con i quali si cerca di dimostrare la non contraddittorietà? e ancor prima, naturalmente, di una *dimostrazione* della stessa?

La non contraddittorietà di una teoria matematica è una affermazione che non riguarda gli enti matematici di cui parla la teoria, ma le dimostrazioni, le formule, il linguaggio insomma in cui è scritta la teoria. Per dimostrarla, occorre che sia una affermazione matematica.

Per trattare la questione, anticipiamo che un modo di procedere è il seguente: trattare il linguaggio come una entità matematica, con la sua complicata struttura statica e dinamica, oggetto di una teoria matematica (metateoria) e trasformare le affermazioni relative al linguaggio e alla teoria in esso formulata (teoria oggetto) in eventuali teoremi della metateoria.

Hilbert considerava i simboli come oggetti concreti e nell'aritmetica elementare le operazioni matematiche si identificavano per lui con le operazioni fisiche: $2 + 3 = 5$ significava che la concatenazione fisica delle stringhe ||| e |||| produce la stringa |||||[1]. Gli enunciati aritmetici elementari (senza quantificatori) erano la rappresentazione o addirittura l'esecuzione grafica di

[1] Le successioni finite di sbarrette sono la rappresentazione unaria dei numeri che sarà adottata nelle macchine di Turing.

tali operazioni. Le derivazioni erano anch'esse catene di liste strutturate di simboli.

Nel 1904, per un frammento molto semplice di logica e aritmetica così concepite - e intrecciate, perché le manipolazioni di simboli coincidevano con quelle aritmetiche - Hilbert aveva elaborato una dimostrazione di non contraddittorietà con un metodo diverso da quello che tutti pensavano fosse l'unico possibile: invece di esibire un modello aveva svolto un ragionamento per induzione sulla lunghezza delle derivazioni. La lunghezza è un parametro numerico, e una misura, assegnabile a oggetti fisici discreti e ordinati, come sono le successioni finite di simboli.

Con poche speciali regole logiche, la dimostrazione riusciva perché doveva considerare derivazioni che erano successioni di formule con un legame interno, molto semplice. La dimostrazione di non contraddittorietà doveva mostrare che nessuna di queste successioni di formule poteva terminare con una contraddizione e il ragionamento, per induzione, si riduceva a considerare l'effetto dell'ultima regola applicata. Ma con sistemi più ricchi di regole, ad esempio il principio di induzione, la complicazione sembrava intrattabile[2].

Una soluzione più soddisfacente e generale consiste nel ricorrere a una codifica esplicita ed esterna dei simboli. Proprio perché non usava una codifica (forse non gli era venuto in mente) ma concepiva direttamente i simboli come oggetti concreti, risultava poi difficile a Hilbert precisare gli strumenti matematici con i quali trattare i linguaggi - a parte l'induzione sulla lunghezza di una successione - al di là della loro descrizione intuitiva come "finitisti".

Nell'epoca dei calcolatori e dei programmi la funzione di una codifica numerica non dovrebbe essere difficile da capire. Tutto quello che si fa e si legge sul calcolatore, parole, disegni, musica, è tutto regolato da programmi che sono scritti, alla base, in binario, quindi in un linguaggio aritmetico.

Quando si scrive in un programma di scrittura che è dotato del controllo dell'ortografia, se si scrive una parola sbagliata all'interno di una frase, ad esempio "l'aritmetica è incompelta", il programma si accorge della presenza di un numero che non rappresenta una parola corretta all'interno del numero che rappresenta la frase, e manda un comando di sottolineare in rosso la corrispondente parola sullo schermo.

[2]Poincaré aveva apprezzato la genialità dell'idea, ma aveva anche profeticamente avvertito che nel caso generale sarebbe occorsa una forma di induzione infinitaria (cioè con formule quantificate) che avrebbe reso circolare il procedimento.

Se si ordina al programma di scrittura di sostituire nel testo "compelta" evidenziata con "completa", il programma soggiacente sostituisce il numero per "compelta", che è in una certa posizione di una più lunga sequenza, con quello di "completa", mettendo poi in moto gli altri programmi che ritrasformano sullo schermo quel numero nella nuova parola. Si sostituisce una stringa di 0 e 1 in una stringa più lunga con un'altra stringa, o un numero con un altro numero in un posto di una espressione. La sintassi è regolata e controllata sui codici numerici degli elementi linguistici. I compilatori trasformano il linguaggio macchina in linguaggi di livello superiore.

Il meccanismo della codifica numerica è alla base della dimostrazione di Gödel. Nella dimostrazione di Gödel c'è una parte *hard* e una parte *soft*, secondo una distinzione spesso usata tra matematici, una parte di calcolo duro e una parte di argomentazione lieve ed elegante.

La parte *hard* riguarda la costruzione della separazione concettuale tra teoria e metateoria e l'uso dell'aritmetica come metateoria formalizzata. La formalizzazione della metateoria non è necessaria solo per poter eseguire dimostrazioni relative ai fatti linguistici, ma si rivela essenziale per sviluppare gli argomenti specifici usati nella dimostrazione, che sono quelli dei paradossi, e costituiscono la parte *soft* della prova di Gödel.

Gödel si ispira a paradossi come quello del mentitore: il paradosso è provocato dalla considerazione di una frase come la seguente centrata

"questa frase è falsa"

che non può essere né vera né falsa, o meglio che è sia vera che falsa.

Gödel riconosce che questo e altri paradossi sono permessi dalla coincidenza di linguaggio e metalinguaggio che si verifica nelle lingue naturali, dove si possono costruire frasi che parlano di frasi della stessa lingua.

Per riproporre questa situazione per l'aritmetica, occorre innanzi tutto che per essa si verifichi la stessa coincidenza di linguaggio e metalinguaggio, e che quindi il metalinguaggio - che di solito è l'italiano - sia aritmetico. Realizzato il contesto generale, occorre poi ovviare al fatto che quasi tutti i paradossi includono parole come "questa" o "io", dette indicali[3], con un riferimento spaziale o temporale, che non sono presenti nel discorso matematico; infine occorre utilizzare un concetto diverso da quello di "verità", perché l'obiettivo non è quello di arrivare a una contraddizione per l'aritmetica, ma

[3]Si potrebbe dire che le virgolette, spesso usate per la menzione di una parola, sono elementi tipici del metalinguaggio e non del linguaggio.

quello di svolgere un ragionamento per assurdo che mostri come una ipotesi porti a una contraddizione, e quindi debba essere negata.

Il compito più impegnativo è il primo, una sorta di costruzione manuale dell'equivalente di un compilatore.

5.1 L'aritmetizzazione dei linguaggi

L'aritmetizzazione[4] è una *codifica* numerica di tutti gli elementi e delle operazioni che si svolgono su di un linguaggio formalizzato.

La condizione che il linguaggio sia formalizzato è essenziale come punto di partenza: non vi devono essere ambiguità o lacune, questa volta per il motivo che la metateoria è matematica. La metateoria in linea di principio costituisce e costruisce il linguaggio oggetto, anche se il modo più naturale di pensare è che il linguaggio sia dato in un metalinguaggio informale come l'italiano e lo si codifichi: ma il linguaggio con i suoi simboli non numerici potrebbe essere solo una rappresentazione di comodo, adatta ai sensi, e la vera sostanza solo i numeri codice.

Attraverso la codifica numerica tutte le affermazioni di carattere linguistico, sintattico, relative alle espressioni della teoria diventano esse stesse espressioni aritmetiche[5]. L'aritmetica diventa la *metateoria sintattica* della *teoria oggetto* studiata[6].

Per l'applicazione che interessa, la teoria oggetto è l'aritmetica. Per fissare le idee, quando si parla di aritmetica si pensi a PA.

Per chiarezza, benché non ci sia bisogno né qui né nel seguito di ricorrere esplicitamente agli assiomi, riportiamo la presentazione formale dell'aritmetica che è nota oggi come *aritmetica di Peano*, o PA.

Il linguaggio contiene l'uguaglianza, la costante $\underline{0}$, il simbolo funzionale a un posto s e i simboli funzionali a due posti[7] $+$ e $\cdot$. I termini

$$\underbrace{s(s(\ldots s}_{n\text{ volte}}(\underline{0})\ldots))$$

[4]Da non confondere con quella dell'analisi realizzata nell'Ottocento.

[5]Per vedere una presentazione dettagliata di questo lavoro, si può consultare il classico E. Nagel e J. R. Newman, *La prova di Gödel*, Boringhieri, Torino, 1961.

[6]La metateoria sintattica potrebbe anche essere una teoria diversa dall'aritmetica, ad esempio la teoria degli insiemi. L'uso di questa come metateoria è preferibile, anzi necessario, se si vuole trattare anche la semantica e non solo la sintassi.

[7]Per essi si usa come tradizione nella scrittura semiformale la notazione infissa, come anche per $=$.

sono i numerali $\underline{n}$, per ogni $n \in \mathbb{N}$.

Gli assiomi sono

$$
\begin{array}{l}
\forall x(\underline{0} \neq s(x)) \\
\forall x \forall y(s(x) = s(y) \to x = y) \\
\forall x(x + 0 = x) \\
\forall x \forall y(x + s(y) = s(x + y)) \\
\forall x(x \cdot \underline{0} = \underline{0}) \\
\forall x \forall y(x \cdot s(y) = x \cdot y + x) \\
\varphi(\underline{0}) \land \forall x(\varphi(x) \to \varphi(s(x))) \to \forall x \varphi(x)
\end{array}
$$

quest'ultimo essendo uno schema, lo schema di induzione, cioè un assioma per ogni formula $\varphi(x)$ del linguaggio.

Gödel non faceva riferimento a questa assiomatizzazione, ma alla parte aritmetica dei *Principia Mathematica* di Russell e Whitehead. PA non è la teoria minimale per cui valgono le dimostrazioni di Gödel, ma quelle più deboli non hanno un significato matematico interessante[8].

L'aritmetizzazione di un linguaggio è lunga e laboriosa perché deve rispettare alcuni requisiti essenziali generali e alcuni specifici per il linguaggio aritmetico in vista della dimostrazione dell'incompletezza.

Non basta assegnare un numero a ogni elemento dell'alfabeto, e poi a ogni lista finita di simboli dell'alfabeto che sono espressioni corrette, e poi a ogni successione finita di espressioni corrette che sono collegate tra loro in modo da formare una derivazione.

Bisogna innanzi tutto che gli insiemi di numeri assegnati a diverse categorie linguistiche siano disgiunti e decidibili, e che i numeri contengano tutte le informazioni relative ai simboli.

Una possibile assegnazione di numeri ai simboli dell'alfabeto si presenta - con possibili varianti inessenziali - nel seguente modo, dove si usa la scomposizione in fattori primi dei numeri per codificare le informazioni: il numero $13^3 \cdot 17^2$ ad esempio codifica il terzo simbolo funzionale (il 3 esponente di 13) che ha due argomenti (il 2 esponente di 17).

[8] Chi è interessato può consultare un manuale come E. Mendelson, *Introduzione alla logica matematica*, Boringhieri, Torino, 1977.

x_1	3
x_2	3^2
x_3	3^3
$\vdots$	$\vdots$
0	5
$=$	$7 \cdot 11^2$
$<$	$7^2 \cdot 11^2$
s	$13 \cdot 17$
$+$	$13^2 \cdot 17^2$
$\cdot$	$13^3 \cdot 17^2$
$\neg$	19
$\wedge$	23
$\vdots$	

A ogni espressione legale, termini e formule, è associato un numero grazie alla possibilità di codificare con numeri le sequenze finite di numeri. Questi numeri sono detti *gödeliani* o *numeri di Gödel* delle entità linguistiche che rappresentano.

Da un numero di questo tipo, appartenente a una particolare categoria, e che *prima bisogna saper riconoscere che è un numero di questo tipo*, si devono saper estrarre i numeri delle espressioni componenti: se è il numero di un termine i numeri degli eventuali sottotermini, se è il numero di una formula i numeri del suo segno logico, delle sottoformule, delle variabili libere, dei termini che vi compaiono. Devono potersi definire diversi predicati e relazioni corrispondenti a quelle sintattiche.

I numeri non possono quindi essere assegnati arbitrariamente alle espressioni, salvo che in certa misura all'alfabeto iniziale, ma devono essere composti con operazioni aritmetiche effettivamente eseguibili ed invertibili sui numeri delle sottoespressioni componenti.

Una realizzazione dell'obiettivo potrebbe consistere nell'assegnare i numeri nel seguente modo:

- se n è il gödeliano di φ

$$2^{19} \cdot 3^n \text{ è il gödeliano di } (\neg\varphi)$$

- se n è il gödeliano di φ e m il gödeliano di ψ,

$$2^{23} \cdot 3^n \cdot 5^m \text{ è il gödeliano di } (\varphi \wedge \psi)$$

- ...
- se φ è $R(t_1, \ldots, t_n)$, dove R è l'i-esimo simbolo relazionale, e $r_1, \ldots, r_n$ sono i gödeliani di $t_1, \ldots, t_n$,

$$2^{7^i \cdot 11^n} \cdot 3^{r_1} \cdot \ldots \cdot p_n^{r_n} \text{ è il gödeliano di } \varphi$$

- ...

In questo modo si evitano anche le parentesi[9].

Alle derivazioni, se $n_1, \ldots, n_r$ sono i gödeliani delle formule che la costituiscono in successione lineare, si può associare il numero $2^{n_1} \cdot \ldots \cdot p_{r-1}^{n_r}$.

A livello sintattico devono essere eseguite varie operazioni a cui non si pensa mai quando si ragiona, come estrarre da (il gödeliano di) una parola[10] il (gödeliano del) simbolo che occorre all'i-esimo posto, riconoscere ad esempio se tale simbolo è una variabile, eseguire sostituzioni in certi posti delle parole, dove ci sono (gödeliani di) variabili libere, e altre operazioni analoghe.

Tutte queste operazioni sintattiche devono trasformarsi in operazioni sui numeri-codice ed essere definibili nella metateoria aritmetica.

Per verificare che sono definibili, Gödel presenta una lista di definizioni precise e dettagliate.

Perché, cosa vuol dire che sono definibili?

Non basta che esista, come esiste, una formula aritmetica $Ter(x)$ tale che

$$\mathbb{N} \models Ter(x)[n] \text{ se e solo se } n \text{ è il gödeliano di un termine}$$

e una $Form(x)$ tale che

[9] Le parentesi sono necessarie nella scrittura usuale, dove una espressione è presentata come una concatenazione di successioni, e quindi come successione; la struttura dei gödeliani proposta sopra corrisponde a concepire le espressioni some alberi, o come successioni di successioni; la posizione prefissa dei simboli di composizione rende superflue le parentesi ai fini del *parsing*.

[10] Useremo questa forma di espressione, "(il gödeliano di) una formula", ad esempio, per abituare progressivamente a fare il passaggio di saltare direttamente alla dizione "una formula", ignorando o cancellando "(il gödeliano di)" quando è ovvio dal contesto che si parla dei codici delle formule.

$$\mathbb{N} \models Form(x)[n] \text{ se e solo se } n \text{ è il gödeliano di una formula}$$

e così via[11].

Occorre come minimo che $\mathbb{N} \models$ sia sostituito da $T \vdash$, dove T è la metateoria.

Se si è definito ad esempio l'insieme degli assiomi della teoria, vuol dire che si è trovata una formula aritmetica $Ass(x)$ per la quale per ogni assioma φ della teoria, se n è il suo gödeliano, $Ass(\underline{n})$ è dimostrabile nella metateoria aritmetica. E viceversa, se si dimostra $Ass(\underline{n})$, se ne possono anche dimostrare alcune conseguenze, ad esempio che $S(\underline{n})$, dove $S(x)$ è la formula che definisce i (gödeliani degli) enunciati, e poi scomponendo n si deve poter risalire al vero enunciato, che deve essere proprio un assioma della teoria, e si deve riconoscere quale, se un assioma per il successore, o per l'uguaglianza, o un caso dell'induzione - e in tal caso per quale formula su quale variabile - o una tautologia logica e di che forma - sempre *dimostrando* qualche formula per $\underline{n}$ che dà tali informazioni.

Tutte queste sono informazioni e operazioni positive, ma ci sono anche quelle negative. Se $Contr(x)$ definisce le contraddizioni ad esempio,

$$Contr(x) \leftrightarrow \exists y(Form(y) \wedge x = 2^{23} \cdot 3^{y} \cdot 5^{2^{19} \cdot 3^{y}}$$

e se n non è (il gödeliano di) una contraddizione, si deve poter dimostrare $\neg Contr(n)$, per poter concludere che una dimostrazione, o le dimostrazioni di un certo tipo, non terminano con una contraddizione.

Analogamente, se n non è un assioma, si deve poter dimostrare $\neg Ass(\underline{n})$, per poter escludere ad esempio che un numero proposto come dimostrazione sia una dimostrazione, se non lo è.

In buona sostanza, le formule che si utilizzano per le definizioni non possono essere qualunque, devono permettere che valga la completezza: che sia sempre dimostrabile o $Ass(\underline{n})$ o $\neg Ass(\underline{n})$, o $Contr(n)$ o $\neg Contr(n)$, per tutti gli enunciati costruiti con formule che definiscono le nozioni rilevanti[12].

L'effettività delle operazioni è riflessa dalla completezza - perché *le operazioni si svolgono nella metateoria, e se questa è una teoria matematica i fatti sono stabiliti da teoremi*.

[11] Questo è quanto si intende di solito con "definibile in una struttura".

[12] Per evitare confusioni con altre nozioni di definibilità meno vincolanti, si parla in questo caso di "rappresentabilità per numerali" (*numeralwise representable*), ma continueremo a parlare di definibilità.

Siccome la completezza della teoria è proprio la questione in discussione, e che risulterà tra l'altro falsa in generale, non la si può assumere, ma la si deve provare per tutte le formule usate per le definizioni; né può essere trattata caso per caso e trovata per tentativi. Le formule trovate da Gödel e usate per l'aritmetizzazione della sintassi soddisfano la condizione: la desiderata completezza vale uniformemente grazie alla forma particolare di tali formule (risultato che potrebbe essere inaspettato), e la loro forma particolare a sua volta dipende dal tipo di definizioni effettive utilizzate, come ora diremo.

5.1.1 Funzioni ricorsive primitive

Il carattere di effettività delle nozioni e operazioni sui gödeliani corrispondenti alle nozioni e operazioni sintattiche viene riassunto in un nuovo concetto, quello delle funzioni numeriche "ricorsive primitive". Per ottenere queste funzioni numeriche, e relazioni connesse[13], occorrono sempre solo alcune mosse definitorie, variamente combinate, a partire da funzioni elementari. Praticamente, come funzioni di partenza, oltre a funzioni che permettono scambio e identificazione di variabili e proiezioni, basta la sola funzione "successore", che genera i numeri stessi ed è rappresentata da un simbolo primitivo del linguaggio; gli schemi definitori necessari sono la composizione e la ricorsione primitiva.

La composizione di funzioni g e $h_1, \ldots, h_r$ è la funzione f che per ogni x

$$f(\overline{y}, x) = g(h_1(\overline{y}, x), \ldots, h_r(\overline{y}, x)),$$

mentre una funzione f si dice definita per ricorsione primitiva da g e h se sono valide le equazioni

$$\begin{cases} f(\overline{y}, 0) & = & g(\overline{y}) \\ f(\overline{y}, x+1) & = & h(\overline{y}, x, f(\overline{y}, x)) \end{cases}$$

dove $\overline{y}$ indica eventuali parametri.

La ricorsione primitiva permette, note g e h e fissati i parametri, di calcolare per ogni n i valori di $f(\overline{y}, n)$ riconducendoli ad argomenti minori, fino a che si arriva a $f(\overline{y}, 0)$; è equivalente all'iterazione, in cui si parte da $f(\overline{y}, 0)$ e per n volte (contatore) si applica la seconda equazione.

La varietà di schemi definitori che si utilizzano è grande, ma tutti si riducono a quelli citati; ad esempio una funzione definita con l'operatore di minimo ristretto, o con una ricerca limitata, tipo

[13]Una relazione è ricorsiva primitiva se la sua funzione caratteristica lo è.

$$f(x, y) = \text{il minimo}\, i < y \,\text{tale che}\, g(x) = 0$$

è ricorsiva primitiva.

Una funzione definita per casi con due condizioni esaustive e mutuamente esclusive che siano ricorsive primitive

$$f(x) \quad = \quad \begin{cases} g(x) & \text{se } P(x) \\ h(x) & \text{se } Q(x) \end{cases}$$

è ricorsiva primitiva.

Poi ci sono tutte le varianti della ricorsione primitiva.

La dimostrazione delle varie riducibilità richiede in generale almeno un lungo capitolo di un trattato sulle funzioni numeriche.

Le operazioni e relazioni della sintassi formalizzata sono tutte ricorsive primitive[14]. Ad esempio l'insieme dei termini ha una definizione di questo genere

$$\begin{aligned} Ter(x) \leftrightarrow x = 2^5 & \\ \vee \ & \exists i < x(x = 2^{3^i}) \\ \vee \ & \exists y, u < x(u = 2^{13 \cdot 17} \wedge Ter(y) \wedge x = 2^u \cdot 3^y) \\ \vee \ & \exists y, z, u < x(u = 2^{13^2 \cdot 17^2} \wedge Ter(y) \wedge Ter(z) \wedge x = 2^u \cdot 3^y \cdot 5^z) \\ \vee \ & \exists y, z, u < x(u = 2^{13^3 \cdot 17^2} \wedge Ter(y) \wedge Ter(z) \wedge x = 2^u \cdot 3^y \cdot 5^z). \end{aligned}$$

Le funzioni ottenute in questo modo sono - dimostra Gödel con una pirotecnica padronanza della teoria dei numeri - definibili, e definibili con i requisiti indicati, di soddisfare le esigenze di completezza per la effettività del doppio movimento della codifica-decodifica.

La mossa cruciale dal punto di vista matematico è quella di trovare una funzione definibile che codifichi mediante numeri le successioni finite arbitrarie di numeri; per mezzo di questa, si riesce a trasformare l'iterazione in una definizione esplicita.

Dal punto di vista logico, Gödel si accorge che l'effettività di calcolo di queste funzioni si traduce nella loro definibilità mediante formule nelle quali tutti i quantificatori sono ristretti, cioè del tipo $\exists x < t\varphi$ o $\forall x < t\varphi$. Allora mentre da una parte le ricerche limitate garantiscono la calcolabilità effettiva, dall'altra queste formule, quando i confini sono numeri, si trasformano in disgiunzioni e congiunzioni.

[14]A dire il vero sono anche meno, rientrano nella classe più ristretta delle funzioni elementari.

Inoltre le funzioni ottenute in questo modo sono definibili in un'aritmetica molto debole, bastano gli assiomi per la somma e per il prodotto[15].

5.2 Paradossi

Se si codifica il linguaggio dell'aritmetica con i numeri, cioè nell'aritmetica stessa, le formule, o almeno alcune di esse, vengono ad avere due significati, uno quello numerico, l'altro quello decodificato. Nel dire di cosa parla una formula si può slittare da un significato all'altro, come se si passasse in una fenditura spazio-temporale, o attraverso uno specchio di Alice mentre si legge la formula stessa.

Ad esempio una formula del tipo

$$\exists y, z, u(x = p_0^y \cdot p_1^z \cdot p_2^u \land y = 5929 \land z = 3 \land u = 9)$$

potrebbe essere letta come una descrizione di x come un numero i cui fattori primi sono 2, 3 e 5 e tale che gli esponenti dei fattori nella sua scomposizione hanno i valori indicati, cioè che

$$x = 2^{5929} \cdot 3^3 \cdot 5^9$$

oppure che x è il gödeliano di $< x_1 x_2$ (ovvero della formula che comunemente si scrive $x_1 < x_2$), essendo

$$x = 2^{7^2 \cdot 11^2} \cdot 3^3 \cdot 5^{3^2}$$

e facendo riferimento alla codifica sopra indicata.

La sovrapposizione di significati matematici e linguistici si avrebbe anche se teoria e metateoria fossero diverse, se la metateoria è matematica. Ma se coincidono, si rende possibile l'autoriferimento, che apre la strada alla parte *soft* della dimostrazione, quella basata sui paradossi e che costituisce anche la parte divertente, e che si fa fatica a convincersi che non sia un trucco di magia.

[15] La codifica usa la potenza e la scomposizione dei numeri in fattori primi; questo metodo richiama alcune intuizioni di Leibniz sulla rappresentazione dell'articolazione dei concetti che Gödel ha ammesso essere state fonte d'ispirazione; ma Gödel ha anche dimostrato che basta la moltiplicazione per codificare con numeri le successioni finite di numeri, usando le congruenze e il teorema cinese del resto.

Nella prima parte del Novecento si è fatto un gran parlare di antinomie e paradossi[16]. Il tutto è iniziato con la scoperta delle contraddizioni nella teoria degli insiemi ingenua, come era usata allora. Uno dei principi che sembravano evidenti, ed uno dei più utili, era il *principio di comprensione*, che postula che data una condizione qualsiasi esiste l'insieme degli enti che soddisfano quella condizione: la libertà di collezionare, o il passaggio dalle proprietà agli insiemi, dall'intensione all'estensione.

Formalizzato, il principio si scrive, se si indica con $\varphi(x)$ una condizione qualunque:

$$\exists X \forall x (x \in X \leftrightarrow \varphi(x))$$

che significa che gli elementi di X sono tutti e soli gli enti che soddisfano φ.

L'antinomia di Russell (1901) si ottiene considerando l'insieme di tutti gli insiemi che non contengono se stessi come elementi. Tale insieme è contraddittorio. La contraddizione si ottiene facilmente dal principio di comprensione sopra scritto, e si vede meglio in formule che a parole, ponendo al posto di $\varphi(x)$ la formula $x \notin x$, leggi "x non appartiene a se stesso", ottenendo quindi un X tale che

$$\forall x (x \in X \leftrightarrow x \notin x);$$

di qui, sostituendo per particolarizzazione X a x,

$$X \in X \leftrightarrow X \notin X.$$

Alcune antinomie di carattere matematico erano state trovate prima, a proposito delle totalità dei cardinali e degli ordinali infiniti. Altre furono trovate subito dopo. Alcune erano di tipo matematico, altre di tipo linguistico.

Kurt Grelling inventò l'antinomia dell'aggettivo "eterologo". Si dica che un aggettivo è autologo se si applica a se stesso, altrimenti lo si dica eterologo. Ad esempio "corto" è corto, mentre "lungo" non è lungo.

[16]Si chiamano *antinomie* le vere e proprie (presunte) contraddizioni; con *paradossi* si intende di solito qualcosa che sembra un'antinomia ma ha o dovrebbe avere una via d'uscita, una spiegazione. Ma la distinzione non è sempre netta: chi propone una antinomia di solito la presenta come una vera e propria antinomia, ma altri la possono vedere in modo diverso. In particolare è discutibile che possa esistere una vera antinomia, cioè una contraddizione assoluta: sono sempre locali, e inducono ad abbandonare o a modificare un particolare progetto.

Se l'aggettivo "eterologo" si applica a se stesso, allora eterologo è autologo, quindi non eterologo, e quindi non si applica a se stesso; se non si applica a se stesso, allora eterologo è eterologo, e quindi si applica a se stesso.

La contraddizione è fondata sulla voluta confusione tra uso e menzione della parola, o tra aggettivo e parola: è la parola "corto" che è corta, non l'aggettivo; l'aggettivo "corto" non è né corto né lungo. Gli aggettivi si applicano alle parole che li esprimono, non agli aggettivi stessi.

Un altro paradosso è espresso da "il minimo numero non definibile con meno di venticinque sillabe", definito dalla formula appena scritta che ha ventiquattro sillabe (attribuito da Russell a G. C. Berry). L'antinomia di Richard consiste nel mostrare che l'insieme dei numeri reali definibili, che dovrebbe essere numerabile, non lo è.

Il pullulare delle antinomie era considerato da molti un argomento a sfavore dei concetti nuovi che si volevano introdurre nella matematica, da quelli insiemistici a quelli logici: non erano matematica.

Finché Gödel non ha eseguito la sua dimostrazione, questo è stato il sentimento più diffuso tra i matematici, di sospetto e fastidio per tali forme di ragionamento. Zermelo cita proprio i paradossi della definibilità tra quelli che dovrebbero essere esclusi dalla sua assiomatizzazione della teoria degli insiemi, nel 1908.

Con Gödel i paradossi vengono ad assumere un carattere positivo, come tecnica dimostrativa. Nell'introduzione al suo lavoro sull'incompletezza, Gödel osserva che "l'analogia di questa argomentazione con l'antinomia di Richard salta agli occhi. Anche con 'il mentitore' sussiste una stretta affinità ... abbiamo una proposizione, la quale afferma la propria indimostrabilità"; in una nota aggiunge: "In generale tutte le antinomie epistemologiche possono essere sfruttate per una simile dimostrazione di indecidibilità".

5.2.1 Il mentitore

L'antinomia del mentitore consiste nell'immaginare una persona che enuncia la frase

"io sto mentendo".

Si chiede allora se tale frase sia vera o falsa: se è vera, allora chi la pronuncia sta mentendo, e quindi ha appeno detto una falsità; se è falsa, chi la pronuncia non sta mentendo, ha detto una verità, e la frase detta è vera.

Un'alternativa si presenta anche con la frase "questa frase è falsa"

o varianti.

La dimostrazione di incompletezza usa come si è detto un artificio analogo, sostituendo "verità" con "dimostrabilità" (in una teoria fissata, di cui si è eseguita l'aritmetizzazione).

L'idea di ispirarsi ai paradossi è venuta a Gödel quando si è scontrato con una variante del mentitore nel tentativo di definire l'insieme delle proposizioni vere dell'analisi, in vista di una dimostrazione di non contraddittorietà relativa. La sua conclusione, poi riscoperta da Alfred Tarski, è stata che l'insieme delle proposizioni vere non è definibile, ammessa la non contraddittorietà della metateoria, perché avrebbe portato a una antinomia del tipo del mentitore; ma si è anche accorto che l'argomento si poteva riadattare con una diversa nozione che fosse definibile.

La difficoltà di trasposizione dell'antinomia dovuta alla presenza di indicali[17] è superata con l'autoriferimento, un meccanismo diventato di moda nel Novecento[18], anche se già noto e utilizzato proficuamente, ad esempio da Saccheri nella sua *Logica*.

L'autoriferimento è la relazione che si stabilisce come per mezzo di uno specchio tra un oggetto e la sua immagine; per realizzarlo per mezzo di frasi prive di indicali occorrono due livelli di discorso, linguistico e metalinguistico, coincidenti.

Gödel ha fatto vedere come costruire un enunciato che, interpretato non come enunciato aritmetico quale è, ma come enunciato metateorico sul linguaggio, si legge "questo enunciato è indimostrabile", o se si vuole

"io sono indimostrabile".

Per fare questo occorre che nel linguaggio aritmetico esistano nomi di enunciati, i gödeliani.

Supponiamo di aver definito la funzione

[17]Se non si vuole ammettere che il paradosso del mentitore mostri la incoerenza della mente umana, si può discutere se "questo momento", che è implicito nella forma "sto mentendo", sia sufficientemente delimitato, o sufficientemente lungo da includere tutta la espressione della frase; e se "questa", o anche altre indicazioni più precise, abbia un riferimento ben definito.

[18]Una ubriacatura di autoriferimento si trova in D. R. Hofstadter, *Gödel, Escher, Bach: un'eterna ghirlanda brillante*, Adelphi, Milano, 1984.

$$sost(x,y) \quad = \quad \begin{cases} \text{(il gödeliano de) la formula che risulta dalla formula} \\ \text{(di gödeliano) } y \text{ sostituendo } \underline{x} \text{ alla sua unica variabile} \\ \text{libera,} \qquad\qquad\qquad\qquad\qquad\qquad \text{se esiste} \\ \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{altrimenti} \end{cases}$$

Supponiamo anche di aver definito $not(x)$ come una funzione che a ogni (gödeliano di una) formula x associa (il gödeliano de) la negazione di x.

Indichiamo con $sostnot(x)$ la composizione $sost(x, not(x))$, cioè la funzione che, data una formula di gödeliano x che abbia una variabile libera, fornisce il gödeliano della formula che si ottiene sostituendo il numerale di x nella negazione della formula di gödeliano x.

Le funzioni in questione sono ricorsive primitive e quindi definibili mediante formule aritmetiche; indichiamo come usa con la sottolineatura (come per i numerali) le formule o i termini che le definiscono.

Con tali premesse, consideriamo la formula con x libera

$$Teor(\underline{sostnot}(x)).$$

Sia n il gödeliano di questa formula. Se sostituiamo il numerale di n alla variabile x nella negazione di questa formula otteniamo

$$\neg Teor(\underline{sostnot}(\underline{n})).$$

Il gödeliano di questo enunciato è, per le proprietà della funzione *sostnot*, proprio $sostnot(n)$. Il trucco è tutto qui.

Se si pensa che il gödeliano sia il nome, ecco realizzato l'"io non sono dimostrabile":

$$sostnot(n) \text{ è il nome di } \neg Teor(\underline{sostnot}(\underline{n})).$$

L'enunciato $\neg Teor(\underline{sostnot}(\underline{n}))$, chiamato enunciato di Gödel, risulta indimostrabile, e la sua negazione pure, sicché si ha l'incompletezza.

L'indecidibilità dell'enunciato di Gödel non si stabilisce tuttavia nello stesso modo immediato della contraddittorietà dell'enunciato del mentitore, in base al significato intuitivo e alla dicotomia vero-falso. Sono richieste alcune ipotesi sulla teoria su cui torneremo.

5.2.2 Richard

L'antinomia di J. Richard (1905) riguarda la definibilità. Sembra naturale pensare che, fissato un qualunque linguaggio, il numero delle definizioni possibili in tale linguaggio sia infinito sì, ma del tipo detto *numerabile*, cioè tale che esse si possano elencare una dopo l'altra, come i numeri naturali[19]. Si pensi all'ordine alfabetico. Richard considera i numeri reali compresi tra 0 e 1, che si possono identificare con le successioni infinite di 0 e 1, in espansione binaria.

Quelli definibili si possono organizzare in una matrice infinita di questo tipo

$$\begin{array}{ccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \dots \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & \dots \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & \dots \\
\vdots & & & & & & & & & &
\end{array}$$

dove ogni riga rappresenta un numero reale definibile.

Si indichi con $r(n,m)$ l'elemento all'incrocio della n-esima riga e dell'm-esima colonna; Richard introduce l'antidiagonale (che in verità è dovuta a Cantor nella sua dimostrazione della non numerabilità dei numeri reali) che è la funzione

$$1 - r(n,n).$$

L'antidiagonale è una successione di 0 e 1, che come si vede subito è diversa da tutte le righe: differisce dalla riga n-esima nel posto n-esimo. Quindi l'antidiagonale individua una nuova successione di 0 e 1 che non è tra quelle definibili.

La definizione dell'antidiagonale tuttavia è appena stata data e sembra corretta e legittima, in quanto eseguita con operazioni aritmetiche sulla matrice. Se l'antidiagonale fosse definita dalla q-esima definizione si avrebbe

$$1 - r(m,m) = r(q,m)$$

per ogni m e una contraddizione per $m = q$.

Se l'antidiagonale di Cantor dimostra la non numerabilità del continuo, l'antidiagonale di Richard che cosa dimostra? Secondo Richard stesso nulla;

[19]Bisogna naturalmente che l'alfabeto sia finito, o al più numerabile.

egli non riteneva la sua una vera contraddizione (un divertimento?), e la risolveva dicendo che la definizione dell'antidiagonale non aveva significato: se fosse stata una riga, la q-esima, al momento di mettere la riga q-esima nella enumerazione per costruire la matrice si sarebbe già dovuto conoscere tutta la matrice[20].

In verità non occorre *conoscere* tutta la matrice, ma sarebbe necessario avere a disposizione un linguaggio che permetta di esprimere sia fatti aritmetici sia fatti relativi alla definibilità aritmetica (e si sappia come farlo). La matrice stessa è individuata utilizzando il concetto di definibilità.

Nell'adattamento dell'argomento di Richard, la matrice di Gödel che porterebbe a una contraddizione sarebbe, fissata una enumerazione di tutte le formule:

$$c(n,m) = \begin{cases} 1 & \text{se la proposizione che si ottiene sostituendo} \\ & \text{il numerale di } m \text{ nella } n\text{−esima formula} \\ & \text{è dimostrabile} \\ 0 & \text{altrimenti} \end{cases}$$

Se questa funzione fosse definibile da una formula aritmetica, lo sarebbe anche l'insieme $\{m \mid c(m,m) = 0\}$, e se la formula che definisce quest'ultimo fosse la q-esima, si avrebbe $c(q,m) = 1$ se e solo se $c(m,m) = 0$, una contraddizione per $m = q$.

Ma Gödel non vuole ottenere una antinomia. In effetti la conclusione di questo argomento è, se si vuole evitare una contraddizione, che una funzione come la $c(n,m)$ che genera la matrice non è definibile in modo da poter essere rappresentata da una formula aritmetica.

Ma perché esattamente non dovrebbe essere definibile, per mezzo della formula $Teor$ prima indicata? Per capirlo, consideriamo prima la formulazione esatta della versione richardiana della dimostrazione di Gödel.

[20]L'osservazione di Richard sarà tuttavia accolta e sviluppata da Russell e Poincaré come una confutazione delle definizioni impredicative; si chiamano così le definizioni con le quali nel definire un ente si fa riferimento a una totalità a cui l'ente stesso risulta a posteriori appartenere. Tali definizioni sono frequenti e innocue nel discorso comune; ad esempio si definisce lo studente migliore di una classe esaminando i voti di tutti gli studenti della classe, incluso quello che risulterà il migliore. Diventano sospette quando si pensa che le definizioni creino gli enti che definiscono. Tale definizioni però sono essenziali per non mutilare parti importanti della matematica, ad esempio intervengono nel teorema di Bolzano-Weierstrass per il calcolo infinitesimale.

La matrice che, se si vuole adottare l'impostazione di Richard, si può associare alla dimostrazione di Gödel è la seguente, dove compare anche il segno $\uparrow$ di "indefinito":

$$g(n,m) = \begin{cases} 1 & \text{se la proposizione che si ottiene sostituendo} \\ & \text{il numerale di } m \text{ nella } n\text{−esima formula} \\ & \text{è dimostrabile} \\ 0 & \text{se la negazione della proposizione che si ottiene sostituendo} \\ & \text{il numerale di } m \text{ nella } n\text{−esima formula} \\ & \text{è dimostrabile} \\ \uparrow & \text{altrimenti} \end{cases}$$

Qui non c'è contraddizione nell'ammettere che l'antidiagonale sia una riga, la q-esima, perché $g(q,q) = 1 - g(q,q)$ non è una contraddizione se $g(q,q) = \uparrow$, e questa risulta in effetti una, e l'unica, via d'uscita: $g(n,m)$ è definibile e q è il gödeliano di una formula tale che $g(q,q)$ è indimostrabile insieme alla sua negazione.

La differenza rispetto alla matrice con $c(n,m)$ è che con $g(n,m)$ il compito della dimostrazione sta proprio nel provare che l'antidiagonale è una riga, trovando la formula che la definisce, e che sarà la q-esima, cosa che Gödel ha esplicitamente trovato.

La matrice $c(n,m)$ era definita da una alternativa tra il caso in cui una formula fosse dimostrabile e il caso in cui non fosse dimostrabile, come se fossero esaustivi, in analogia al carattere bivalente della nozione di verità. La matrice $g(n,m)$ distingue tra il caso in cui una formula è dimostrabile, il caso in cui la negazione della formula è dimostrabile e un terzo in cui ipoteticamente non si dà nessuno dei due casi precedenti.

Il caso $\uparrow$ non va preso come un terzo valore (che si potrebbe scrivere ad esempio 2), ma semplicemente come eventualità non presa in considerazione, quindi non menzionata; naturalmente si ha un'aporia, una funzione parziale, ma mentre nella definizione di $c(n,m)$ compare $\neg Teor$, nella definizione di $g(n,m)$ compare solo $Teor$.

Ebbene, "dimostrabile" è definibile con alcune caratteristiche di effettività richieste dalla trattazione, anche se non la completezza, precisamente si ha che se n è il gödeliano di un teorema, allora $Teor(n)$ è dimostrabile. Invece il "non dimostrabile" di $c(n,m)$ non ha questa proprietà: non necessariamente se n è il gödeliano di un enunciato non dimostrabile si dimostra $\neg\, Teor(\underline{n})$.

Se si mette una negazione davanti alla formula che definisce "dimostrabile", l'effetto è dirompente. Il motivo è che queste formule non sono ristrette.

$Teor(x)$ è una formula del tipo $\exists y Dim(y, x)$, dove $Dim(y, x)$ è la formula che definisce una relazione ricorsiva primitiva "y è (il gödeliano di) una dimostrazione della formula (di gödeliano) x".

Se m è (il gödeliano di) una dimostrazione della formula (il cui gödeliano è) n, allora $Dim(\underline{m}, \underline{n})$ è un teorema, e anche quindi $Teor(\underline{n})$.

$\neg\, Teor(\underline{n})$ è della forma $\forall y \neg Dim(y, \underline{n})$.

Se n è il gödeliano di un enunciato non dimostrabile, allora nessun m è (il gödeliano di) una dimostrazione di n, e si dimostrano tutti i $\neg Dim(\underline{m}, \underline{n})$, ma non è detto che si dimostri $\forall y \neg Dim(y, \underline{n})$.

Si vede qui l'importanza del fatto che le funzioni ricorsive primitive siano definibili con formule ristrette, per le quali vale la completezza, mentre anche un solo quantificatore universale non ristretto la fa saltare[21].

Nel caso di $c(n, m)$, l'insieme $\{m \mid c(m, m) = 0\}$ non è l'insieme $\{m \mid \neg Teor(sost(\underline{m}, \underline{m})\}$, né definibile con altra formula.

Questa decisiva minuzia, gravida di conseguenze, non si poteva certo scoprire ragionando superficialmente sul fatto che l'enunciato di Gödel afferma la propria indimostrabilità o con tutte le altre versioni intuitive.

5.3 I teoremi di Gödel

L'enunciato di Gödel è stato originariamente dimostrato indecibile nel modo seguente.

Supponiamo che $\neg\, Teor(\underline{sostnot}(\underline{n}))$, di gödeliano $sostnot(n)$, sia dimostrabile e sia m (il gödeliano di) una sua dimostrazione; allora sarebbe dimostrabile

$$Dim(\underline{m}, \underline{sostnot}(\underline{n}))$$

e

$$\exists y Dim(y, \underline{sostnot}(\underline{n}))$$

cioè

$$Teor(\underline{sostnot}(\underline{n})).$$

[21] Per evitare troppi tecnicismi, non siamo del tutto rigorosi. Per la precisione, l'affermazione è vera per le formule, mentre se un insieme è definibile sia da una formula $\exists y \varphi(y, x)$ sia da una formula $\forall y \psi(y, x)$, φ e ψ ristrette, allora l'insieme è ancora decidibile.

Se la teoria non è contraddittoria, questo non è possibile.

Ma neanche la sua negazione $Teor(\underline{sostnot}(\underline{n}))$ è dimostrabile, se si fanno ipotesi ragionevoli. Poiché si è visto che $\neg\, Teor(\underline{sostnot}(\underline{n}))$ non è dimostrabile, nessun numero è (il gödeliano di) una sua dimostrazione, quindi si dimostra

$$\begin{gathered}\neg Dim(\underline{1}, \underline{sostnot}(\underline{n}))\\ \neg Dim(\underline{2}, \underline{sostnot}(\underline{n}))\\ \neg Dim(\underline{3}, \underline{sostnot}(\underline{n}))\\ \vdots\end{gathered}$$

e così via per ogni m

$$\neg Dim(\underline{m}, \underline{sostnot}(\underline{n})).$$

Se fosse dimostrabile $Teor(\underline{sostnot}(\underline{n}))$ si avrebbe però anche una dimostrazione di $\exists y Dim(y, \underline{sostnot}(\underline{n}))$.

Gödel chiama ω-coerenza la proprietà di una teoria di non ammettere che per alcuna formula si dimostri sia $\varphi(\underline{n})$ per ogni numerale, sia $\exists x \neg\varphi(x)$. Si conclude che se la teoria è ω-coerente allora $Teor(\underline{sostnot}(\underline{n}))$ non è dimostrabile[22].

La ω-coerenza è una proprietà simile alla non contraddittorietà, ma più forte; vuole esprimere indirettamente il fatto che esiste un modello, quello standard, perché gli elementi del modello sono tutti descrivibili dai numerali. In seguito J. B. Rosser dimostrò anche questa metà del risultato assumento solo la non contraddittorietà.

Una teoria formulata in un linguaggio aritmetico si dice "abbastanza potente" se è in grado di sviluppare al suo interno la parte di aritmetica necessaria per l'aritmetizzazione, cioè addizione e moltiplicazione. La teoria si dice "effettivamente assiomatizzabile" se l'insieme dei (gödeliani dei) suoi assiomi è ricorsivo (cioè definibile come abbiamo sopra richiesto per Ass, con anche $\neg Ass(\underline{n})$ dimostrabile per gli n che non sono assiomi).

Allora,

Teorema 1. *Se T è una teoria effettivamente assiomatizzabile, abbastanza potente e non contraddittoria, allora T è incompleta.* □

[22] $sostnot(n)$ è un esempio di un numero m che non è (il gödeliano di) un teorema ma per cui $\neg Teor(\underline{m})$ non è dimostrabile.

La costruzione dell'enunciato di Gödel è effettiva ed uniforme, cioè si ottiene sempre allo stesso modo a partire dagli assiomi e dalle regole[23] di T.

Infatti se G è l'enunciato indecidibile di Gödel per T, allora sia T con l'aggiunta di G come assioma, sia T con l'aggiunta di $\neg G$ sono non contraddittorie. Tuttavia l'esistenza e la costruzione dell'enunciato indecidibile si ripete automaticamente nello stesso modo per queste nuove teorie[24].

Il secondo teorema di incompletezza afferma che l'enunciato che esprime la non contraddittorietà di T non è dimostrabile in T.

Un enunciato che esprima la non contraddittorietà di T si ottiene in diversi modi, come, se $Teor_T$ è la formula che definisce i teoremi di T:

$$NCon_T \leftrightarrow \neg\exists x(Teor_T(x) \wedge Teor_T(\underline{neg}(x)))$$

oppure

$$NCon_T \leftrightarrow \neg\forall x Teor_T(x)$$

o ancora in altre versioni; se ad esempio c è il gödeliano di $0 = 1$

$$NCon_T \leftrightarrow \neg Teor_T(\underline{c}).$$

Teorema 2. *Nelle stesse ipotesi*[25] *su T del teorema 1, $NCon_T$ non è dimostrabile in T.* □

La dimostrazione richiede un'aritmetizzazione della dimostrazione del primo teorema di incompletezza, che si è presentata in modo informale; si era osservato che se la teoria non è contraddittoria allora $\neg Teor_T(\underline{sostnot}(\underline{n}))$ non è dimostrabile; ricordando che $sostnot(n)$ è il gödeliano di $\neg Teor_T(\underline{sostnot}(\underline{n}))$, aritmetizzando l'argomento che giustificava quell'osservazione, e che è una dimostrazione, si avrebbe

$$T \vdash NCon_T \rightarrow \neg\, Teor_T(\underline{sostnot}(\underline{n})),$$

[23]Le regole logiche possono essere diverse, ma devono poter permettere un certo numero e tipo di inferenze che si possono precisare.

[24]Cambia un po' $Teor$ al variare di Ass, ma mantiene la stessa struttura.

[25]Tuttavia "abbastanza potente" deve qui essere inteso in modo più forte. I particolari sono stati definitivamente precisati da S. Feferman nel 1960.

e se si dimostrasse $NCon_T$ si dimostrerebbe $\neg\, Teor_T(\underline{sostnot}(\underline{n}))$.

Questa dimostrazione, molto laboriosa, fu data solo nel 1939 da David Hilbert e Paul Bernays e si trova nelle loro *Grundlagen der Mathematik*[26].

In una teoria più forte dell'aritmetica, ad esempio la teoria degli insiemi con l'assioma dell'infinito, non c'è alcuna difficoltà a dimostrare la non contraddittorietà di PA; è interessante vedere cosa è il minimo che si debba aggiungere a PA per ottenere tale risultato: ad esempio forme di induzione transfinita su ordinali, o linguaggi con variabili anche per insiemi[27] ma con assiomi deboli, o logiche diverse. Gödel ha osservato che "considerando la situazione da un punto di vista puramente matematico, le dimostrazioni di non contraddittorietà sulla base di assunzioni metamatematiche più forti opportunamente scelte (come sono state date da Gerhard Gentzen e altri) sono altrettanto interessanti e foniscono importanti intuizioni sulle strutture teorico-dimostrative della matematica".

5.4 Funzioni ricorsive

All'indagine logica restava aperto un problema importante, quello di sapere se esistesse un metodo effettivo di decisione per la logica del primo ordine. Per l'aritmetica la dimostrazione di Gödel provava che non c'era, o lo faceva sospettare se si accetta un legame tra metodi effettivi e funzioni rappresentabili nell'aritmetica. La funzione caratteristica di $\{x \in \mathbb{N} \mid Teor(x)\}$, cioè la funzione f tale che per ogni x

$$f(x) = \begin{cases} 1 & \text{se } Teor(x) \\ 0 & \text{altrimenti} \end{cases}$$

non è rappresentabile nell'aritmetica, altrimenti l'antidiagonale $1 - c(n, n)$ della matrice $c(n, m)$ sarebbe una riga, e si avrebbe una contraddizione, come abbiamo già visto.

Per completare l'argomento occorre precisare il concetto di metodo effettivo di decisione in modo che diventi equivalente a una definizione logica o matematica. Infatti per provare che un problema è decidibile occorre e basta esibire un algoritmo per risolvere il problema, con una dimostrazione

[26] Lo stesso Gödel fece tuttavia in seguito osservare che la conclusione del secondo teorema è sensibile alla forma dell'enunciato $NCon_T$, e infatti per certe formulazioni della non contraddittorietà si è riusciti in seguito a dimostrarla, in T.

[27] Linguaggi del secondo ordine.

di correttezza (cioè che l'algoritmo fa quello che deve). Se invece si sospetta che un metodo di decisione non esista, e si vuole provare a dimostrarlo, il concetto stesso di metodo effettivo deve essere trattabile come un oggetto matematico.

Gödel ha esteso la classe delle funzioni ricorsive primitive, che erano più che sufficienti per l'aritmetizzazione, ottenendo le funzioni ricorsive con l'aggiunta dell'operatore di minimo alla composizione e all'iterazione.

Si dice che una funzione f è definita con l'operatore di minimo o per minimizzazione a partire da g se

$$f(\overline{y}, x) \simeq \mu z(g(\overline{y}, z) = 0)$$

che si legge: $f(\overline{y}, x)$ è il più piccolo numero z tale che $g(\overline{y}, z) = 0$, se un tale numero esiste[28]. Con $\simeq$ indichiamo che o entrambi i membri sono definiti, e allora sono uguali, oppure entrambi sono indefiniti.

Le funzioni che si ottengono possono non essere ovunque definite e si chiamano ricorsive parziali.

Le funzioni ricorsive parziali costituiscono una classe di funzioni aritmetiche; ma definire il concetto di "calcolo" è un'altra questione, ed esistono diverse alternative che risultano tutte equivalenti, nel senso che le funzioni che risultano calcolabili secondo le varie definizioni coincidono con le funzioni ricorsive parziali.

Secondo Gödel, le equazioni che definivano le sue funzioni non dovevano essere intese come sistemi di equazioni funzionali (in una variabile funzionale) che garantivano solo l'esistenza della funzione soddisfacente il sistema, ma come assiomi per derivare i valori delle funzioni in una logica dotata solo di regole per il trattamento delle equazioni (precisate da S. C. Kleene, 1934-36); il calcolo è quindi una derivazione da equazioni.

Secondo il λ-calcolo di Alonzo Church (1936), i valori dovevano essere ottenuti in una teoria dedicata alla pura nozione di funzionalità, con il concetto primitivo di applicazione di una funzione a una funzione. Il calcolo è un processo di riduzione di termini, con opportune regole di riduzione, a una forma normale che è quella dei numerali.

Alan Turing (1936) ha inventato un modello astratto di macchina, definendo calcolabili le funzioni calcolabili da una macchina di Turing e calcoli i processi eseguiti dalle macchine.

[28] $g(\overline{y}, z) = 0$ è un modo di standardizzare una condizione ricorsiva.

Gödel ha dichiarato che solo l'analisi delle procedure meccaniche svolta da Turing lo convinse che si fosse ottenuta una nozione rigorosa adeguata del concetto di procedimento effettivo.

Il motivo per cui Gödel riteneva importante la chiarificazione di tale nozione era che per mezzo di essa si poteva anche dare la versione più generale possibile dei teoremi di incompletezza.

Già era ovvio che essi si riferivano non solo all'aritmetica ma anche a ogni *verwändter Systeme* (come indicato nel titolo del suo storico lavoro); e nell'enunciato dei teoremi presentati in **5.3** si parla di "teorie" qualsiasi soddisfacenti le ipotesi enunciate. Ma ora si poteva parlare in generale di "sistemi formali", definiti non necessariamente con assiomi e regole in linguaggi adatti per l'aritmetica, ma per mezzo dell'idea della produzione meccanica dei teoremi. Un sistema formale è un insieme di regole per trasformare parole date in altre parole, dove le regole possono essere applicate da un agente umano o incorporate ed eseguite da una macchina, e non sono necessariamente le regole deduttive.

I sistemi formali e le tecniche usate per costruirli, per usarli e per studiarli diventano omogenee, si tratta sempre di macchine. Scompare la necessità dell'aritmetizzazione e il fenomeno dell'incompletezza si può davvero concepire come presente in qualsiasi metodo gli esseri umani possano adottare per la generazione della matematica.

L'incompletezza si riferisce dunque non a particolari prodotti dell'attività matematica, ma all'attività matematica stessa. Nello stesso tempo essa diventa imparentata con altri fenomeni che riguardano in particolare il comportamento delle macchine.

5.5 Macchine di Turing

Una macchina di Turing (MT) è un insieme finito e non contradditorio di istruzioni; l'alfabeto in cui sono scritte le istruzioni è composto da due insiemi finiti non vuoti disgiunti, quello dei simboli di lettura e scrittura $\{\alpha_1, \ldots, \alpha_m\}$ e quello degli stati $\{q_0, \ldots, q_n\}$, e dalle lettere D e S. Tra i simboli si include sempre un simbolo speciale $*$ (*blank*).

Usiamo le lettere greche α, β come metavariabili sui simboli, e M come metavariabile su $\{D, S\}$.

Un'istruzione è una quadrupla[29] di uno di questi due tipi:

[29] In alcune presentazioni, le istruzioni delle MT invece che quadruple sono quintuple.

$$q_i \, \alpha \, \beta \, q_j$$
$$q_i \, \alpha \, M \, q_j$$

e un insieme di istruzioni si dice non contraddittorio se non contiene due istruzioni diverse che iniziano con la stessa coppia $q_i \, \alpha$.

L'effetto della macchina è quello di trasformare parole di un tipo particolare in parole dello stesso tipo; si tratta di parole in cui compare un solo stato e gli altri elementi sono simboli, come in

$$\alpha_{i_1}, \ldots, \alpha_{i_j}, q_h, \alpha_{i_{j+1}}, \ldots, \alpha_{i_r}.$$

Un'interpretazione intuitiva fisicamente figurata è quella di pensare a un nastro potenzialmente infinito (estendibile quanto si vuole se necessario), diviso in caselle, in cui ogni casella contiene un simbolo; una testina di lettura e scrittura in ogni istante è posizionata su di una casella - nell'esempio della parola di sopra la casella su cui è scritto $\alpha_{i_{j+1}}$ e la macchina si trova in un particolare stato q_h. Si può pensare al nastro come nastro infinito, ma tale che in ogni istante solo un numero finito di caselle non siano vuote. Una casella vuota si può pensare marcata con *blank*.

La macchina fisica, oltre a una testina che le permette di leggere (nel senso di reagire in modo diverso a simboli diversi) e scrivere, ha un corpo che è una scatola nera, di cui si sa solo che può assumere un numero finito di configurazioni o stati distinti, in funzione della storia precedente di attività svolte e successione di stati assunti. La storia è riassunta a tutti gli effetti dall'ultimo stato assunto[30].

L'effetto di una istruzione $q_i \, \alpha \, \beta \, q_j$ è il seguente: se la macchina in un dato istante si trova nello stato q_i e (è posizionata su una casella su cui) legge α, allora scrive al suo posto β e all'istante successivo passa allo stato q_j; l'effetto di una istruzione $q_i \, \alpha \, M \, q_j$ è che se la macchina in un dato istante si trova nello stato q_i e legge α, lo lascia scritto e all'istante successivo si muove di una casella a destra o a sinistra a seconda che M sia D o S.

Se la macchina in un dato istante si trova nello stato q_i e legge α, e tra le sue istruzioni non ce ne è nessuna che inizi con la coppia $q_i \, \alpha$, la macchina si ferma. La macchina si ferma anche se, data la situazione descritta, esiste un'istruzione[31] $q_i \, \alpha \, \alpha \, q_i$.

Come alfabeto di simboli per i calcoli numerici è sufficiente l'insieme di due simboli, la sbarretta $|$ e $*$. Questa affermazione è però un difficile risultato della

[30] Anche se questo, con un numero finito di stati e storie arbitrariamente lunghe, non è del tutto evidente.

[31] Nelle istruzioni q_i non è necessariamente diverso da q_j come α non necessariamente diverso da β.

teoria delle MT. Di solito è più comodo avere a disposizione altri simboli, in particolare ad esempio simboli che servono a delimitare il campo finito significativo, ed estendibile, del nastro entro cui si lavora.

Diverse convenzioni standardizzano la scrittura delle macchine di Turing:

- si conviene che ci sia sempre almeno una istruzione che inizia con q_0, che è detto *stato iniziale*, e che la macchina all'istante iniziale sia nello stato q_0 e la testina sia posizionata sul primo simbolo non *blank* a sinistra;

- per l'arresto abbiamo già detto quale soluzione adottiamo, anche se sono possibili altre soluzioni, come richiedere che la macchina si fermi sempre in un particolare stato finale q_H, H per *halt*, oppure ammettere anche insiemi contraddittori con la condizione che la macchina si ferma se ha due istruzioni contrastanti;

- per la rappresentazione dei numeri interi positivi, se i simboli sono $\{|, *\}$, il numero n è rappresentato in input da una successione ininterrotta di $n+1$ sbarrette $|$, in particolare lo 0 da $|$;

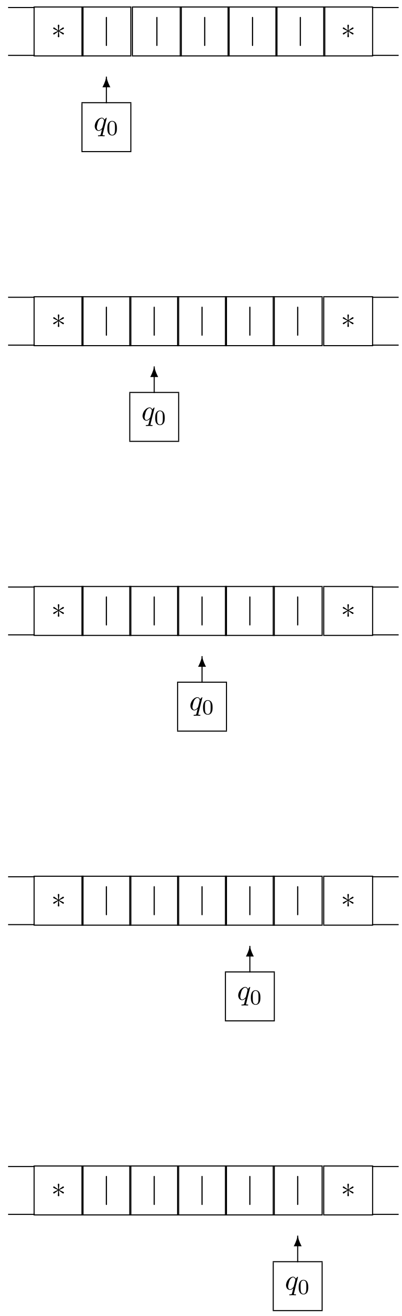
- una coppia di numeri $\langle m, n\rangle$ sarà rappresentata in input da $|||\cdots| * ||\cdots|$, con $m+1$ sbarrette seguite da $*$ e da $n+1$ sbarrette;

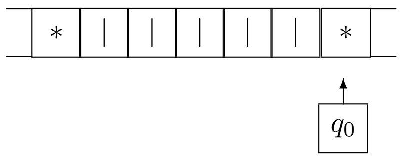
- quando la macchina si ferma, per l'interpretazione del risultato numerico sono possibili diverse alternative, che possono variare anche da macchina a macchina, ma devono essere dichiarate - e la loro adozione influenza la scrittura della macchina: si può pretendere che il numero sia rappresentato come in input da una parola $||\cdots|$ e che queste siano le uniche caselle non *blank* del nastro; oppure che ci possano anche essere altre caselle non *blank* ma che la testina sia posizionata su una delle caselle del blocco di sbarrette consecutive che rappresentano il risultato, e all'inizio o alla fine o in altre posizione intermedie; oppure si può convenire che quando la macchina si ferma un opportuno decodificatore conti il numero totale finito di $|$ sul nastro (per questo è necessario che sia ben delimitata da simboli speciali la porzione di nastro da controllare), e altre soluzioni sono ancora possibili.

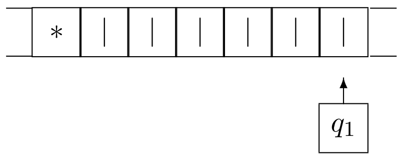
Esempio: Le seguenti istruzioni

$$q_0 \mid D\, q_0$$
$$q_0\ * \mid q_1$$

costituiscono una MT che computa il successore di ogni numero; la macchina si ferma nello stato q_1 posizionata sulla $|$ più a destra del risultato. La tipica successione dei nastri e degli stati è la seguente:







Se si aggiungono

$$q_1 \mid S\, q_1$$
$$q_1 * D\, q_2$$

si ottiene un'altra macchina che computa la stessa funzione, ma si ferma nello stato q_2 posizionata sulla prima $\mid$ di sinistra del risultato.

Osservazione 1. *Esistono macchine diverse che eseguono lo stesso compito.*

Poiché una MT *è* l'insieme delle sue istruzioni, le due macchine di sopra sono chiaramente diverse, ma si possono chiamare equivalenti; è anche facile immaginare che ne esistono infinite diverse tra di loro ed equivalenti (basta aggiungere istruzioni inoffensive, con stati mai raggiunti).

Esempio: La seguente macchina calcola la differenza $m-n$, secondo la seguente semplice strategia iterativa: per ogni sbarretta nel secondo numero, si sposta nel primo e ne cancella una. Il secondo argomento serve da *contatore* del numero di iterazioni necessarie. Complicazioni sono dovute al dover tener conto della grandezza relativa dei due argomenti, al fatto che solo n sbarrette delle $n+1$ che rappresentano n devono contare per l'iterazione, e al dover mantenere invariata la parte centrale della parola input per ritrovare la stessa situazione a ogni ciclo (quindi la $\mid$ cancellata ogni volta da m è la prima a sinistra).

$$q_0 \mid D\, q_0$$
$$q_0 * D\, q_1$$
$$q_1 \mid D\, q_2$$

$$
\begin{array}{l}
q_2 \mid D\, q_3 \\
q_3 \mid D\, q_3 \\
q_3 * S\, q_4 \\
q_4 \mid\ * q_4 \\
q_4 * S\, q_5 \\
q_5 \mid S\, q_5 \\
q_5 * S\, q_6 \\
q_6 \mid S\, q_6 \\
q_6 * D\, q_7 \\
q_7 \mid\ * q_7 \\
q_7 * D\, q_0 \\
q_2 * S\, q_8 \\
q_8 \mid\ * q_8 \\
q_8 * S\, q_9 \\
q_9 * S\, q_{10}
\end{array}
$$

Il primo gruppo di istruzioni

$$
\begin{array}{l}
q_0 \mid D\, q_0 \\
q_0 * D\, q_1 \\
q_1 \mid D\, q_2 \\
q_2 \mid D\, q_3 \\
q_3 \mid D\, q_3 \\
q_3 * S\, q_4
\end{array}
$$

porta la testina alla fine dei due gruppi di sbarrette, e la successiva istruzione cancella l'ultima. Quindi la macchina torna all'inizio nello stato q_7 con

$$
\begin{array}{l}
q_4 * S\, q_5 \\
q_5 \mid S\, q_5 \\
q_5 * Sq_6 \\
q_6 \mid S\, q_6 \\
q_6 * D\, q_7
\end{array}
$$

e con le due successive istruzioni cancella la prima sbarretta e si sposta a destra riposizionandosi su $\mid$ nello stato q_0 pronta a iterare.

Il gruppo finale di istruzioni

$$
\begin{array}{l}
q_2 * S\, q_8 \\
q_8 \mid\ * q_8 \\
q_8 * S\, q_9 \\
q_9 * S\, q_{10}
\end{array}
$$

entra in funzione quando a destra c'è inizialmente (perché n è 0) o è rimasta una sola sbarretta, che viene semplicemente cancellata e la macchina si riposiziona sul risultato.

Quando $m \geq n$, la macchina dà come risultato $m-n$ e si ferma nello stato q_{10}, posizionata su una sbarretta del risultato. Se $m < n$ si danno due casi, $m+1 = n$ e $m+1 < n$, distinguibili dallo stato finale (esercizio).

Esercizio: Progettare una macchina che dati due numeri m ed n rappresentati da $|\,|\,|\cdots|\ *\,|\,|\cdots|$ decide quale è il maggiore posizionandosi su (una sbarretta di) quello maggiore, o sul *blank* intermedio se sono uguali.

Esercizio: Usando la macchina precedente, disegnare una macchina che per ogni m ed n calcola $m-n$ se $m \geq n$ ed $n-m$ se $n \geq m$, cioè calcola $|m-n|$.

Esempio: Macchina che esegue la somma di due numeri:

$$
\begin{array}{l}
q_0 \mid \ast\ q_1 \\
q_1\ \ast\ D\, q_2 \\
q_2 \mid D\, q_3 \\
q_3 \mid D\, q_3 \\
q_3\ \ast\ \mid q_4 \\
q_4 \mid S\, q_4 \\
q_4\ \ast\ D\, q_5 \\
q_5 \mid \ast\ q_5
\end{array}
$$

Dopo aver cancellato la prima $|$ la macchina si sposta a destra e se trova subito un $\ast$ si ferma nello stato q_2, perché la somma proposta era $0+m$. Altrimenti continua a spostarsi a destra fino al primo $\ast$, lo sostituisce con $|$, torna a sinistra alla prima $|$ e la sostituisce con $\ast$. Il risultato è in ogni caso il numero somma dei due dati.

Tuttavia questa macchina non è elegante, perché è basata su un ragionamento esterno, del programmatore, sul numero totale di sbarrette; le operazioni si calcolano invece di solito con l'iterazione, come nel caso della macchina per la sottrazione, e in quest'altra per l'addizione:

$$
\begin{array}{l}
q_0 \mid \ast\ q_1 \\
q_1\ \ast\ D\, q_2 \\
q_2 \mid \ast\ q_3 \\
q_3\ \ast\ D\, q_4 \\
q_4 \mid D\, q_4 \\
q_4\ \ast\ D\, q_5 \\
q_5 \mid D\, q_5 \\
q_5\ \ast\ \mid q_6 \\
q_6 \mid S\, q_7 \\
q_7 \mid S\, q_7 \\
q_7\ \ast\ S\, q_8 \\
q_8 \mid S\, q_9
\end{array}
$$

$q_9 \mid S\, q_9$
$q_9 \;*\; D\, q_2$

Dagli esempi dovrebbe essere chiaro come si può realizzare la ricorsione primitiva con le istruzioni di una MT. Inoltre due macchine si possono combinare in modo da ottenere una macchina che calcola dapprima la funzione calcolata dalla prima, quindi sul risultato la funzione calcolata dalla seconda, cioè la composizione delle due funzioni; la composizione delle macchine si ottiene semplicemente unendo i due insiemi di istruzioni dopo aver traslato gli indici degli stati della seconda macchina in modo che non interferiscano con quelli della prima, il q_0 della seconda diventando il q_h finale della prima (dopo aver eventualmente modificato la prima in modo da garantire che si fermi su uno stato di indice massimo, e posizionata sulla prima sbarretta a sinistra del risultato).

Esempio: La macchina *Copy*, che dato un numero ne scrive uno uguale a partire da due caselle più a destra (della fine del primo). Usiamo un simbolo ausiliario \$ che mettiamo ripetutamente su ogni sbarretta del numero, e in corrispondenza al quale andiamo a scrivere una sbarretta al fondo della parola. Alla fine, se si vogliono avere il numero originale e la copia, i \$ andranno risostituiti con |.

$q_0 \mid \$\, q_0$
$q_0\, \$\, D\, q_1$
$q_1 \mid D\, q_1$
$q_1 * D\, q_2$
$q_2 * D\, q_3$
$q_3 * \mid q_5$
$q_3 \mid D\, q_4$
$q_4 \mid D\, q_4$
$q_4 * \mid q_5$
$q_5 \mid S\, q_5$
$q_5 * S\, q_6$
$q_6 * S\, q_7$
$q_7 \mid S\, q_8$
$q_8 \mid S\, q_8$
$q_8\, \$\, D\, q_0$
$q_7\, \$\, D\, q_9$
$q_9 \;* D\, q_9$

(Mancano le istruzioni per ripristinare le sbarrette al posto dei dollari).

In modo analogo si può scrivere una MT che copia una parola qualunque da un'altra parte desiderata, ma precisata, del nastro.

La macchina *Copy* è necessaria quando si distrugge uno degli input nel corso del calcolo, ma lo si deve conservare, o per il risultato, o per altri calcoli, oppure lo si vuole duplicare.

Osservazione 2. *Esistono macchine che per certi input non si fermano.*

Ad esempio la macchina data dalle istruzioni

$$\begin{array}{l} q_0 \mid Dq_1 \\ q_1 \mid Dq_0 \\ q_0 * Dq_0 \end{array}$$

si ferma nello stato q_1 se il numero n in input è pari, e il risultato è n stesso, mentre continua a spostarsi a destra nello stato q_0 se il numero è dispari.

Esercizio: Scrivere una MT che come nel caso precedente si ferma se e solo se n è pari, ma in tal caso dà come risultato 1.

Esercizio: Scrivere una MT che per ogni numero n dà come risultato 1 se n è pari, e 0 se n è dispari.

Esempio: Macchina per dividere per 2.

$q_0 \mid D\, q_1$	
$q_1 \mid D\, q_2$	ci sono almeno due |
$q_2 \mid D\, q_2$	scavalca input
$q_2\ * D\, q_7$	
$q_7 \mid D\, q_7$	scavalca risultato parziale
$q_7\ * \mid q_8$	scrive |
$q_8 \mid S\, q_8$	torna a sinistra
$q_8\ * S\, q_9$	
$q_9 \mid S\, q_9$	
$q_9 * D\, q_{10}$	arrivata all'inizio
$q_{10} \mid\ * q_{10}$	cancella due |
$q_{10} * D\, q_{11}$	
$q_{11} \mid\ * q_{11}$	
$q_{11} * D\, q_0$	itera
$q_0 * \$\, q_0$	input numero dispari
$q_1 * \mid q_3$	rimasta una | nell'input
$q_3 \mid S\, q_4$	
$q_4 \mid\ * q_5$	
$q_5 * D\, q_6$	

Con le ultime istruzioni a partire da $q_1 * \mid q_3$ si sposta la sbarretta residua a fianco del risultato e si posiziona su di esso (funzionano anche se l'input era 0, come il risultato).

$q_0 * \$ q_0$ fa sì che la macchina si fermi con sul simbolo \$ se l'input era un numero dispari; informa che la divisione non è eseguibile.

Con $q_0 * S\, q_{11}$ al posto di $q_0 * \$\, q_0$ la macchina non si fermerebbe.

Invece di $q_0 * \$\, q_0$, si possono dare le seguenti istruzioni, che fanno scrivere a destra il resto 1, cioè $||$.

$$
\begin{array}{l}
q_0 * D q_{12} \\
q_{12} |\, D\, q_{12} \\
q_{12} * D\, q_{13} \\
q_{13} * |\, q_{13} \\
q_{13}\, |\, D\, q_{14} \\
q_{14} * |\, q_{14}
\end{array}
$$

Gli esempi proposti mostrano come tra i compiti che possono risolvere le macchine di Turing ci sia, oltre a quello di calcolare le funzioni, anche quello di riconoscere l'appartenenza o meno a insiemi o relazioni.

Definizione 1. *Si dice che una funzione $f : \mathbb{N} \longrightarrow \mathbb{N}$ è calcolata da una MT M se per ogni input n la macchina M si ferma e dà come risultato il numero $f(n)$. Una funzione f si dice* effettivamente calcolabile *se esiste una MT che la calcola.*

Si dice che una relazione $R \subseteq \mathbf{N}^n$ è effettivamente calcolabile se la sua funzione caratteristica è effettivamente calcolabile.

Data una macchina M e un input n (per semplicità di notazione consideriamo solo questo caso), o meglio un input che è la codifica scelta per n, indicheremo con $M(n)$ il (numero che decodifica il) risultato scritto sul nastro quando M si ferma e quando esso è un risultato numerico secondo le convenzioni di decodifica. Per tenere conto dei casi in cui non si ferma, o il risultato non è numerico, introduciamo il simbolo per "indefinito" $\uparrow$ e scriviamo $M(n)\uparrow$ per dire che M su n non dà un risultato numerico o non dà nessun risultato perché non si ferma.

Per le funzioni non ovunque definite si ha

Definizione 2. *Si dice che una funzione parziale $f \subseteq \mathbb{N}\times\mathbb{N}$ è calcolata da una MT M se per ogni input n $M(n) \simeq f(n)$. Una funzione parziale f si dice* effettivamente calcolabile *se esiste una MT che la calcola.*

Che la definizione di calcolabilità debba ritenersi adeguata è il contenuto della cosiddetta

Tesi di Church Le funzioni che sono calcolabili mediante un procedimento intuitivamente effettivo coincidono con le funzioni effettivamente calcolabili, cioè calcolabili mediante macchine di Turing.

5.6 Problemi indecidibili

La complicazione della definizione di $\uparrow$ con l'alternativa del risultato non numerico o della non terminazione è dovuta al fatto che, quando una funzione non ha un valore definito, come abbiamo visto nel caso della sottrazione, la macchina può segnalare la non esistenza del valore in vari modi, dopo un tempo finito. Resta aperto il problema se possa sempre essere così, cioè sia sempre possibile per funzioni parziali avere una macchina che o calcola il valore, se c'è, oppure ci segnala con qualche artificio che il valore non c'è.

Se una funzione parziale è effettivamente calcolabile, quindi esiste una macchina che permette di ottenere i suoi valori (quando ci sono), è sempre il caso che il dominio di questa funzione è decidibile? Il dominio è dato dagli argomenti sui quali la macchina si ferma con un risultato numerico. Se è decidibile, come nel caso della sottrazione o della divisione, si può estendere la funzione ai casi in cui non è definita ponendo dei valori convenzionali. La divisione può venir trasformata ad esempio in una funzione che dà sempre come risultato la coppia $\langle$quoziente, resto$\rangle$.

Per fare questo, l'importante è che la macchina si fermi, poi guardando il nastro si vede se l'output è un numero o no. Quindi il problema della estensione a una funzione totale è collegato al seguente: data una macchina, si può decidere a priori quali sono gli input su cui si ferma? "A priori" vuol dire non ricorrendo all'ovvia soluzione di far partire la macchina e stare a vedere quando si ferma: se infatti siamo su un caso in cui non si ferma, aspettiamo le calende greche senza avere una risposta.

Si potrebbe pensare che data la semplicità di funzionamento di ciascuna MT, e il fatto che per scriverla bisogna pensare a come si comporta sui vari input possibili significativi (0 o non zero, pari o dispari e cosí via), la cosa sia possibile, ma non è così.

Decidere il problema della fermata (o dell'arresto, o dell'*halt*) significa deciderlo in senso tecnico, cioè mediante una macchina di Turing.

Il problema è simile a quello incontrato per le dimostrazioni di non contraddittorietà, o di altra proprietà linguistica, un problema innanzi tutto di espressione.

Sui nastri delle macchine di Turing si possono scrivere parole che codificano anche alfabeti non numerici, in particolare le istruzioni di una macchina di Turing.

Basta avere due simboli a e b per rappresentare, ad esempio con a, aa, aaa e cosí via, gli stati $q_0, q_1, \ldots$ e con $b, bb, bbb, \ldots$ i simboli $\alpha_1, \alpha_2, \ldots$; data una macchina particolare, diciamo con tre stati e due simboli, si potrà usare $a**$, $aa*$ e aaa per i tre stati e $b*$, bb per i due simboli, oltre a dd per D e ss per S, in modo da avere una lunghezza fissa delle parole che codificano le istruzioni; con altri opportuni simboli di separazione si potrà rappresentare ogni quadrupla della macchina e quindi tutte le istruzioni.

Da un'altra parte del nastro si può codificare con questi stessi simboli un input della macchina rappresentata, quindi eseguire sulle codifiche le varie operazioni di confronto, copia, modifica che farebbe la macchina data su quell'input.

Se indichiamo sempre con M la codifica su nastro delle istruzioni di una macchina M, e se scriviamo $M(n) \Downarrow$ se la macchina M messa in funzione sul nastro con input n si ferma e $M(n) \Uparrow$ se non si ferma, il problema dell'*halt* consiste nel chiedersi se esista una macchina di Turing H tale che per ogni M ed n

$$H(M,n) = \begin{cases} 1 & \text{se } M(n) \Downarrow \\ 0 & \text{se } M(n) \Uparrow \end{cases}$$

Decidere il problema dell'arresto non deve necessariamente implicare di essere in grado di eseguire i calcoli che ogni macchina esegue. Anzi H non dovrebbe proprio comportarsi in questo modo, altrimenti quando replica un calcolo senza terminazione non termina neanche lei. Si può pensare di poter trovare un modo di analizzare le istruzioni che permetta di scoprire se è possibile che si instauri un ciclo, magari classificando i possibili gruppi di istruzioni "pericolosi", in quanto possono produrre cicli, ed esaminandoli appositamente.

Ma è chiaro che il solo pensare a una macchina che soddisfi le condizioni di sopra per H suggerisce di chiedersi se esista una macchina capace di replicare i calcoli di ogni altra macchina. Questo sarebbe già un primo passo nella direzione della plausibilità di H.

Una macchina di Turing si può pensare in due modi: o che sia fisicamente realizzata in modo da comportarsi per ogni input secondo quanto le istruzioni prevedono, e quindi calcolare solo una funzione[32], oppure come istruzioni (programma) da consegnare ad un interprete perché questo le esegua.

Quando, data una macchina, si mostra come la macchina si comporta su determinati input (come negli esempi di sopra), una persona sta interpre-

[32]Le prime macchine calcolatrici costruite da Pascal e Leibniz eseguivano solo una o due operazioni.

tando le istruzioni; chi le interpreta e descrive la successione conseguente di nastri e stati segue, o applica, le istruzioni. Nel compiere il lavoro il soggetto umano mette in gioco solo abilità di lettura e scrittura, oltre alla comprensione di cosa vuol dire fare certi movimenti, come spostare l'occhio o l'indice (testina) a destra e a sinistra, e la compulsione decisa a farlo di fronte alla lettura delle istruzioni. Le abilità richieste e manifeste sono le stesse che hanno portato a definire il concetto di Macchina di Turing. Questo significa che nell'interpretare una MT il soggetto si comporta come una macchina, mette in gioco solo le abilità che definiscono una MT. Lo stesso soggetto, con gli stessi comportamenti, può interpretare ogni MT ed eseguire ogni calcolo di ogni MT. Questa osservazione rende plausibile che esista una macchina unica capace di fare la stessa cosa. Si tratta di una *macchina di Turing universale*.

Una macchina di Turing universale U è una MT tale che innanzi tutto nel suo linguaggio è possibile codificare ogni possibile insieme di istruzioni per una macchina di Turing; quindi per ogni M ed ogni input adeguato a M, indicata con $U(M,n)$ la macchina U messa in azione su un input costituito dalla codifica della macchina M e di un input n per M, si ha

$$U(M,n) \simeq M(n).$$

Teorema 3. *Esiste una Macchina di Turing Universale.* $\Box$

Teorema 4. *Non esiste una macchina di Turing H tale che per ogni MT M ed ogni n*

$$H(M,n) = \begin{cases} 1 & \textit{se } M(n) \Downarrow \\ 0 & \textit{se } M(n) \Uparrow \end{cases}$$

Dimostrazione La dimostrazione si ispira alle antinomie, sfruttando un autoriferimento. L'assunzione da cui bisogna partire è che esiste una enumerazione effettiva di tutte le macchine di Turing,

$$M_0, M_1, \ldots, M_n, \ldots$$

vale a dire che esiste una MT la quale, per ogni n, stampa le istruzioni di M_n. Una tale enumerazione esiste perché gli insiemi finiti di istruzioni possono essere enumerati per numero crescente di istruzioni e stati coinvolti e in un ordine alfabetico. Il fatto che si considerano macchine non ovunque definite facilita il compito, perché non c'è da verificare nulla, se non la correttezza sintattica.

Supponendo dunque per assurdo che il problema dell'arresto fosse decidibile, esisterebbe una macchina H, composta servendosi di quella che da n calcola una rappresentazione di M_n, tale che per ogni i ed n

$$H(i,n) = \begin{cases} 1 & \text{se } M_i(n) \Downarrow \\ 0 & \text{se } M_i(n) \Uparrow, \end{cases}$$

e ne esisterebbe allora una H' tale che per ogni i

$$H'(i) = \begin{cases} 1 & \text{se } M_i(i) \Downarrow \\ 0 & \text{se } M_i(i) \Uparrow. \end{cases}$$

Per costruire H' da H basta predisporre una macchina che dato i lo duplica, e poi inizia a lavorare come H sulla coppia $\langle i, i \rangle$.

Data H', esiste allora H'' tale che per ogni i

$$H''(i) = \begin{cases} 0 & \text{se } M_i(i) \Uparrow \\ \Uparrow & \text{se } M_i(i) \Downarrow. \end{cases}$$

Per avere H'' basta aggiungere ad H' istruzioni che quando questa si ferma fanno un test sul risultato, e se questo è 0 non fanno nulla, se invece è 1 mettono in funzione un ciclo, dato ad esempio con una coppia

$$\begin{array}{l} q_r \mid Dq_s \\ q_s \mid Sq_r. \end{array}$$

H'' è una MT, quindi occorre nella enumerazione, ad esempio è M_h. Ma ora è facile verificare che:

$$M_h(h) \Downarrow \text{ se e solo se } M_h(h) \Uparrow$$

un assurdo, che si può evitare solo negando l'ipotesi che esista la MT H. □

Il contrasto tra questi due risultati, l'esistenza di una macchina universale e l'indecidibilità del problema dell'arresto, e tra le loro dimostrazioni, simili ma con un colpo di coda finale diverso, è sempre fonte di meraviglia; il commento di John von Neumann (1949) è il seguente:

> Si può costruire un automa [macchina] che può fare qualunque cosa fa un altro automa, ma non si può costruire un automa che predica il comportamento di un automa arbitrario. In altre parole, si può costruire un organo che fa qualsiasi cosa che può essere fatta, ma non uno che dica se può essere fatta.

Una semplice riformulazione del teorema è il

Corollario 1. *Il problema dell'arresto per MT è indecidibile.* □

Riferito a un problema, "indecidibile" è diverso da "indecidibile" riferito a un enunciato, come quello di Gödel. Un *problema* non è mai una domanda singola, ma un insieme, o una proprietà, e la soluzione del problema è un metodo che si applica uniformemente e dà la risposta (appartenenza sì o no) per ogni caso possibile. Che un enunciato sia indecidibile in una teoria vuol solo dire che né l'enunciato né la sua negazione sono dimostrabili nella teoria, e basta.

Tuttavia esiste un collegamento, che si vede dalla dimostrazione; quando un problema risulta indecidibile è perché per ogni proposto metodo risolutivo esiste un caso particolare in cui il metodo si trova bloccato e non può dare nessuna risposta, né positiva né negativa; su quel caso particolare, la risposta si comporta come un enunciato indecidible in una teoria che riguardi le macchine di Turing.

6 La metamatematica

Le derivazioni sono oggetti finiti, successioni finite o altre strutture combinatorie[1], e il rapporto tra deduzioni e semantica, o tra ragionamento e verità, può essere impostato in modo rigoroso, addirittura matematico.

La logica non è più soltanto un insieme di regole d'uso, una logica pratica (*logica utens*), ma diventa una teoria (*logica docens*). Questa è stata l'idea di Hilbert quando ha inventato la metamatematica e ha formulato il suo programma:

> La questione della non contraddittorietà per i numeri naturali e per gli insiemi non è una questione isolata, ma appartiene a un grande ambito di questioni gnoseologiche fra le più difficili aventi tonalità specificamente matematiche: al fine di caratterizzare brevemente questo ambito di questioni, cito la questione della *risolubilità* in linea di principio *di ogni problema matematico*, la questione della *controllabilità* a posteriori del risultato di una ricerca matematica, e inoltre la questione relativa ad un *criterio di semplicità* per le dimostrazioni matematiche, la questione del rapporto tra *contenutisticità* e *formalismo* in matematica e in logica, e infine la questione della *decidibilità* di un problema matematico mediante un numero finito di operazioni[2].

La metamatematica è indipendente tuttavia dal programma di Hilbert; è proseguita dopo la risposta negativa fornita dai teoremi di incompletezza, ed era peraltro iniziata prima con alcuni risultati preliminari, innanzi tutto di metalogica.

6.1 Completezza

Per lo sviluppo della teoria logica è essenziale stabilire il rapporto tra la nozione di conseguenza semantica e quella deduttiva, e auspicabilmente che sia di coincidenza. Non è sorprendente né difficile dimostrare che

$$\text{Se } T \vdash \varphi \text{ allora } \varphi \text{ è conseguenza logica di } T,$$

[1] Alberi ad esempio; se concepite come successioni, come minimo possono essere arricchite da rimandi interni (*links*), come negli esempi visti.

[2] D. Hilbert, "Il pensiero assiomatico" (1917), in *Ricerche sui Fondamenti della Matematica* (a cura di V. M. Abrusci), Bibliopolis, Napoli, 1978, pp. 177-88.

che si chiama correttezza o validità (*soundness*) del sistema logico. Viceversa

$$\text{Se } \varphi \text{ è conseguenza logica di } T \text{ allora } T \vdash \varphi$$

è la proprietà di completezza del sistema logico, ed è il teorema di completezza di Gödel. Il problema era stato posto nel libro di Hilbert e Ackermann del 1928 che è stato il primo manuale di logica matematica[3], e Gödel alla ricerca di un argomento di tesi lo risolse nel 1929.

Lo si dimostra in generale dimostrando prima

$$\text{Se } T \text{ è non contraddittoria, allora esiste un modello di } T.$$

ma le due formulazioni sono equivalenti grazie al fatto che da una parte, solo in base alle definizioni

$$T \cup \{\neg\varphi\} \text{ ha un modello se e solo se } \varphi \text{ non è conseguenza logica di } T$$

e dall'altra

$$T \cup \{\neg\varphi\} \text{ è non contraddittoria se e solo se } T \nvdash \varphi$$

con facili passaggi deduttivi.

Si ricordi che la non contraddittorietà è definita in termini della nozione sintattica $\vdash$.

L'esistenza del modello non richiede interventi creativi soprannaturali esterni, ma è guidata dalle leggi interne di funzionamento del formalismo dato (in questo caso i legami logici), come capita probabilmente in tutte le forme di creazione: gli elementi della struttura $\mathcal{M}$ sono i termini del linguaggio, le relazioni sono definite in modo che $R^{\mathcal{M}}$ vale per $t_1, \ldots, t_n$ se $T \vdash R(t_1, \ldots, t_n)$, e così via[4].

Non diversamente in matematica, quando ad esempio si vuole immergere un dominio di integrità in un campo gli elementi dell'estensione sono le frazioni a coefficienti nel dominio.

[3]D. Hilbert e W. Ackermann, *Grundzüge der theoretischen Logik*, Springer, Berli9n, 1928.

[4]Stiamo semplificando al prezzo di notevoli imprecisioni, ma non c'è tempo per maggiori dettagli.

6.2 Semidecidibilità

Mentre la relazione di conseguenza logica è altamente non costruttiva - infinitistica - la relazione di derivabilità si presta a un trattamento più concreto. Per determinare se $T \vdash \varphi$, un procedimento che si può immaginare, grossolano ma effettivo, è quello di elencare tutte le possibili derivazioni che hanno assunzioni prese da T, in ordine sostanzialmente di lunghezza. Se se ne trova una che termina con φ, si ha la risposta positiva.

Il procedimento è molto poco efficiente, perché l'albero delle possibili derivazioni è molto largo: in ogni nodo molte sono le diverse regole che si possono applicare per la prosecuzione della derivazione, e inoltre se T è infinito bisogna prevedere la progressiva aggiunta periodica di un nuovo elemento di T (T lo si suppone enumerato, se infinito, in modo effettivo).

Invece di generare tutte le derivazioni si può anche generare, che è più facile, tutte le successioni finite di formule, quindi o contemporaneamente passarle in rassegna escludendo quelle che non sono derivazioni.

La ricerca (*search*) può essere organizzata in modo che sia esaustiva, nel senso che se la derivazione c'è prima o poi la si trova. Se non esiste tuttavia si rischia di aspettare all'infinito.

Un simile procedimento - che dà sempre la risposta giusta, in un numero finito di passi, per i casi con un tipo di risposta, mentre non dà alcuna risposta per almeno alcuni casi del secondo tipo - si chiama procedimento di decisione parziale, e un problema che ammette un procedimento di decisione parziale di chiama semidecidibile.

Per la logica ne esistono anche di più efficienti della ricerca esaustiva. Ci riferiamo sempre a metodi con i quali il ritrovamento o la costruzione della eventuale derivazione non richieda inventiva o fantasia ma segua regole meccaniche.

Dato φ, se uno vuole costruire una derivazione di φ da T ci sono varie considerazioni e strategie che si possono seguire, anche soltanto considerando la forma sintattica di φ, per non parlare delle conoscenze generali che si hanno su T. Se è una congiunzione $\varphi_1 \land \varphi_2$ si possono considerare separatamente φ_1 e φ_2, dal momento che $T \vdash \varphi_1 \land \varphi_2$ se e solo se $T \vdash \varphi_1$ e $T \vdash \varphi_2$. Se φ è una negazione $\neg\psi$, si può pensare di cercare una dimostrazione per assurdo a partire da ψ, e così via. Queste e simili euristiche possono essere organizzate in una ricerca che pur seguendo regole meccaniche assomiglia a forme di ragionamento reali, e che ha successo, quando la derivazione esiste[5].

[5]Per uno di questi metodi, quello degli alberi di refutazione, se si è interessati si vedano

Grazie all'equivalenza stabilita dal teorema di completezza si può affermare che la relazione di conseguenza logica è semidecidibile, come anche quindi la nozione di validità logica.

Quando un problema è semidecidbile, perché è noto un metodo di decisione parziale, nulla esclude che si riesca a trovare un altro procedimento che è un metodo di decisione: un algoritmo che per ogni caso del problema risponde sempre sì o no, in modo corretto, dopo un numero finito di passi.

La logica predicativa è soltanto semidecidibile, non decidibile, come stabilisce il teorema di Church, 1936, che presentiamo come Corollario al Teorema 4 di **5.6**.

Corollario 2 (Church, 1936)**.** *Il problema della validità logica per linguaggi predicativi è indecidibile.*

Dimostrazione Indichiamo solo i passi fondamentali. Il problema dell'arresto viene *ridotto* al problema della validità logica, o della conseguenza logica.

a) Le funzioni effettivamente calcolabili coincidono con le funzioni ricorsive parziali.

b) Le funzioni ricorsive parziali coincidono con quelle che sono definibili da particolari sistemi di equazioni, nel senso che i loro valori si ottengono per mezzo di derivazioni, a partire dalle equazioni definitorie e dalle proprietà dell'uguaglianza.

c) Quindi si può associare in modo effettivo a ogni calcolo di una macchina di Turing (per una funzione numerica, ricorsiva parzale) una derivazione da un sistema di equazioni (formule).

Grazie a questa corrispondenza, si vede che se il problema della conseguenza logica, e quindi della derivabilità - grazie alla completezza - fosse decidibile, anche il problema dell'arresto lo sarebbe. □

Si dice in breve: *la logica predicativa è indecidibile*, solo semidecidibile.

Si vede facilmente che un insieme è decidibile se e solo se sia l'insieme sia il suo complemento sono semidecidibili: si mettono in funzione, per un n qualunque dato, sia il metodo di decisione parziale per l'insieme, sia quello per il complemento, e dopo un numero finito di passi da una della due macchine si ha la risposta.

Si usano anche i termini "ricorsivamente enumerabile", o "generabile", per "semidecidibile". Il motivo sta nel fatto che, mettendo in funzione un procedimento

---

le dispense di G. Lolli per il corso di Logica Matematica nel sito, o altri manuali.

di decisione parziale, si possono ottenere via via gli elementi dell'insieme, "sputati" per cosí dire dalla macchina man mano che questa si accorge che un numero appartiene all'insieme.

Così si vede anche in che senso una teoria può essere rappresentata come una macchina: il processo di generazione sistematico ed esaustivo di tutte le derivazioni, e delle loro conclusioni, è lo stesso di quello con cui data una funzione ricorsiva parziale f si genera il suo dominio con il metodo detto di *dove-tailing*: la macchina M che calcola f viene fatta lavorare (da un'altra macchina) in questo modo: M inizia a fare un certo numero k di passi del calcolo di $f(0)$, quindi k passi del calcolo di $f(1)$, indi riprende ulteriori k passi del calcolo di $f(0)$, k di $f(1)$ e nuovi k passi di $f(2)$, e così via. Quando trova che il calcolo di $f(i)$ è terminato, ne dà il valore.

Con alcune restrizioni sui linguaggi, si può fare di meglio della semidecidibilità. Abbiamo visto che esiste un metodo di decisione per la validità dei sillogismi. Più in generale, se il linguaggio ha solo predicati a un posto, allora la relazione di conseguenza logica, per tutte le formule, non solo per quelle che corrispondono alle proposizioni categoriche dei sillogismi, è decidibile. Anche la logica proposizionale, con solo i connettivi, è decidibile.

Il problema della decidibilità viene in seguito indagato anche per le teorie matematiche.

6.3 Teorie decidibili

Se una teoria è completa, allora è decidibile: basta di nuovo il procedimento di generazione sistematica di tutte le derivazioni: per ogni φ, dopo un numero finito di passi si incontra o una derivazione di φ o una derivazione di $\neg\varphi$.

Si conferma l'intuizione di Poincaré che un elenco completo degli assiomi della geometria avrebbe portato a meccanizzare la ricerca dei teoremi. Ma il procedimento di generazione sistematico è così inefficiente che chi è interessato al risultato non può esimersi dal cercarne di più efficienti.

Una teoria decidibile, e completa, è quella dei corpi reali ordinati chiusi: questa teoria aggiunge agli assiomi dei corpi ordinati gli infiniti assiomi che affermano che ogni equazione di grado dispari ha almeno una radice.

Il risultato, dovuto a Tarski, si basa su una estensione del teorema di Sturm, a sua volta estensione della regola di Cartesio; le affermazioni di esistenza di una radice diventano equivalenti a una condizione algebrica, e il quantificatore esistenziale scompare. Si tratta poi di far vedere che tutti i quantificatori possono essere eliminati da qualsiasi formula, e che quelle prive di quantificatori sono decidibili.

Il risultato è importante anche per la geometria, in quanto espressa analiticamente.

Per quel che riguarda la praticità, è stato dimostrato tuttavia che ogni possibile algoritmo di decisione deve avere una complessità almeno superesponenziale. Lo stesso vale per altre teorie che sono state dimostrate complete e decidibili, quali la teoria dei campi algebricamente chiusi di caratteristica p fissata (p un primo o 0), la teoria delle algebre di Boole atomiche infinite.

Sono indecidibili invece ad esempio le teorie dei gruppi, degli anelli, dei corpi, dei reticoli. Ma è decidibile la teoria dei gruppi abeliani. Il panorama è molto intricato.

6.4 Modelli non standard

Il teorema di completezza inoltre ha diverse conseguenze importanti dal punto di vista matematico; è uno strumento potente per la costruzione di strutture in algebra, per il fatto che la coincidenza delle nozioni semantiche e sintattiche trasferisce alle prime - alle strutture - le caratteristiche di finitezza delle seconde.

Siccome è evidente che in una derivazione di φ da T compaiono solo un numero finito di elementi di T, perché per definizione una derivazione è una successione finita, è immediato che

T è non contraddittoria se e solo se ogni $S \subseteq T$ finito è non contraddittorio,

e da questa condizione e dalla completezza segue che

Se ogni sottoinsieme finito di T ha un modello, T ha un modello,

che si chiama teorema di compattezza[6].

Teorie composte da un numero infinito di assiomi sono di uso frequente, più di quanto non si pensi se non si riflette sul linguaggio; l'aritmetica di Peano è una di queste, la teoria dei corpi reali chiusi, o quella dei campi di caratteristica zero sono altri esempi. Ma se ne possono costruire anche appositamente per la soluzione di certi problemi.

[6]Ha in effetti legami con la compattezza topologica.

Una applicazione importante, come chiarificazione concettuale e come base per la costruzione dell'analisi degli infinitesimi, è quella dei modelli non standard dell'aritmetica.

Al linguaggio dell'aritmetica, quello di PA o altro, si aggiunga un nuovo simbolo c di costante individuale, e si consideri il seguente insieme di enunciati

$$c \neq \underline{0}, c \neq \underline{1}, \ldots, c \neq \underline{n}, \ldots$$

da aggiungere come assiomi agli altri. Si ha un insieme infinito T con la proprietà che ogni suo sottoinsieme finito è non contraddittorio, ammesso che gli assiomi dell'aritmetica siano non contraddittori. Noi supponiamo che $\mathbb{N}$ sia noto e sia per noi l'insieme dei numeri naturali, che consideriamo il modello standard dell'aritmetica[7]. Ogni sottoinsieme finito degli assiomi di T contiene alcuni degli assiomi della teoria e alcuni dei nuovi enunciati aggiunti. Esso può essere soddisfatto in $\mathbb{N}$ se a c si assegna come interpretazione un numero che sia maggiore di tutti gli n i cui numerali $\underline{n}$ compaiono nell'insieme scelto.

Per il teorema di completezza, l'insieme T ha un modello $\mathcal{M}$, che soddisfa tutti gli assiomi dell'aritmetica ma non può essere $\mathbb{N}$ perché in esso deve esistere un elemento diverso da tutti gli $n \in \mathbb{N}$. $\mathcal{M}$ non è solo diverso, ma non isomorfo a $\mathbb{N}$, anche se numerabile, perché un eventuale isomorfismo non potrebbe assegnare un corrispondente a $c^{\mathcal{M}}$. Un tale modello si chiama modello non standard del'aritmetica[8].

T può essere qualunque, perché il teorema di completezza non è ristretto alle teorie effettivamente assiomatizzabili, quindi T potrebbe essere l'insieme di tutti gli enunciato veri in $\mathbb{N}$. Questo insieme non è effettivo; anzi ha un alto grado di non costruttività, ma può essere preso in considerazione. Segue allora con il ragionamento di sopra che esistono modelli non isomorfi a $\mathbb{N}$ ma che soddisfano esattamente gli stessi enunciati - sono linguisticamente indistinguibili - a conferma che le nozioni di categoricità e completezza divergono.

Si vede dunque che il teorema di completezza di Gödel da una parte dà un fondamento scientifico alla credenza di Hilbert che la non contraddittorietà di un sistema di assiomi implichi l'esistenza di enti soddisfacenti quegli

[7]D'altra parte lo usiamo per la costruzione e la descrizione del linguaggio, ad esempio nella definizione dei numerali di sopra.

[8]La sua struttura è molto complessa e interessante, ma non ci concerne qui.

assiomi; dall'altra lo stesso teorema sanziona la caratteristica del metodo assiomatico (positiva per i sostenitori, negativa per i critici) della pluralità dei modelli. Nello stesso tempo esclude le interpretazioni difensive del metodo che speravano di conservare la categoricità, in quanto permette di dimostrare che nessuna teoria formulata in un linguaggio del primo ordine e che abbia un modello infinito è categorica[9].

6.5 Teorie algebriche

Un modello non archimedeo dei campi si ottiene aggiungendo agli assiomi la lista infinita di enunciati

$$0 < \underline{n} \cdot c < 1$$

per ogni $n \in \mathbb{N}$, c nuova costante, dove $\underline{n} \cdot c$ sta per $\underbrace{c + \ldots + c}_{n \text{ volte}}$.

Spesso molte condizioni che sono formulate senza pensare al linguaggio provocano la introduzione di un numero infinito di assiomi, in particolare se fanno riferimento alla nozione di numero naturale. La negazione dell'assioma di Archimede dovrebbe essere, sopra,

$$\forall n(0 < \underline{n} \cdot c < 1)$$

intendendo

$$\forall n \in \mathbb{N}(0 < \underline{n} \cdot c < 1)$$

ma $\mathbb{N}$ non è definibile nel linguaggio e nella teoria dei campi. Ogni singolo numero invece lo è (se la caratteristica è 0) e si ovvia nel modo visto.

Perché $\mathbb{N}$ non è definibile nella teoria dei corpi, anche in quelli di caratteristica 0? Una risposta che valga in generale non è facile. Se una teoria è completa, in essa non si può definire (un insieme isomorfo a) $\mathbb{N}$, altrimenti relativizzata a questa definizione si potrebbe ripetere la dimostrazione di Gödel.

L'insieme dei termini $\underbrace{1 + \ldots + 1}_{n \text{ volte}}$ è definito nella metateoria, che tuttavia in questi casi non coincide con la teoria. La teoria oggetto potrebbe essere non archimedea.

[9] Per le strutture finite si può invece dare una loro descrizione completa: ad esempio $a \neq b \wedge b \neq c \wedge c \neq a \wedge \forall x(x = a \vee x = b \vee x = c)$ obbliga i modelli ad avere tre elementi.

Una situazione analoga si ha con la nozione di caratteristica 0, o con quella di torsione nei gruppi.

La caratteristica 0 è espressa dall'insieme infinito di assiomi

$$\underbrace{1 + \ldots + 1}_{n \text{ volte}} \neq 0$$

per ogni $n \in \mathbb{N}$.

Si ha il seguente risultato, come conseguenza del teorema di compattezza: se un enunciato vale in tutti i campi di caratteristica 0, esiste un n_0 tale che l'enunciato vale in tutti i campi di caratteristica $> n_0$.

Si sono sviluppati metodi e risultati ulteriori per lo studio delle teorie e dei loro modelli, cioè di classi di strutture che sono i modelli di una teoria, anche al di là degli iniziali teoremi di completezza e di compattezza. L'argomento è il soggetto di una disciplina diventata quasi autonoma, che si chiama appunto "teoria dei modelli"[10].

[10]Sembra quasi superfluo avvertire che l'uso della parola "modello" in logica va nella direzione opposta a quello della modellistica matematica: qui un modello è uno schema formale della realtà, in logica è l'interpretazione (concreta?) di uno schema formale. Tuttavia una conciliazione è possibile, se si interpretano i modelli logici come rappresentazione semplificata, e pur sempre matematica, di un frammento di realtà, mediata o guidata da una formulazione linguistica che compie l'astrazione. Lo schema formale in sé va inteso come una traccia per la costruzione del modello. Quando si fa un modello in scala ad esempio, la teoria è l'insieme dei rapporti di riduzione, e il modello soddisfa i requisiti formali.