DALLA MATEMATICA DEI GRECI AGLI INSIEMI NUMERICI

Tentativi di fondare la matematica

VOL. I

Giangiacomo Gerla[1]

ggerla@unisa.it

[1]Professore Emerito di Matematica dell'Università di Salerno.

INDICE

Introduzione

CAPITOLO 1
LA MATEMATICA PRESSO I GRECI

CAPITOLO 2
CRISI DELLA GEOMETRIA EUCLIDEA

CAPITOLO 3
DAI NUMERI AL CONTINUO GEOMETRICO

CAPITOLO 4
DEFINIRE LA NOZIONE DI ALGORITMO

CAPITOLO 5
DEFINIRE LA NOZIONE DI PROBABILITA'

APPENDICE
NOZIONI BASE E VARIE

INTRODUZIONE

Questo è il primo di due volumetti che raccolgono i miei appunti per i corsi di Matematiche Complementari e che si propongono di introdurre il lettore alle problematiche relative ai fondamenti della matematica. Essi espongono i vari tentativi fatti da matematici e filosofi per fornire basi sicure alla matematica e per capirne i "principii primi" (ammesso che alla matematica si possa dare una base definitiva ed ammesso che per essa esistano principii primi). Poiché gli appunti sono scritti per studenti di un corso della laurea di Matematica, spesso si troveranno definizioni precise ed un po' pedanti e, principalmente, dimostrazioni di teoremi. Tuttavia credo che, almeno questo primo volume, possa essere letto da tutti poiché esiste l'antico diritto di chi legge di saltare pagine, dimostrazioni e parti noiose.
Ovviamente si pone la questione:

esiste veramente l'esigenza di un tale tipo di riflessione sulla matematica ?

Dopotutto sembra non esistere niente di più semplice e sicuro di nozioni come quelle di numero, punto, retta. Purtroppo semplicità e sicurezza sono illusioni come mostrano i vari paradossi emersi nel corso dell'evoluzione della matematica. Il fatto è che siamo tanto abituati a manipolare i concetti matematici che tendiamo a confondere la familiarità che abbiamo acquisito con la conoscenza di tali concetti. Un po' avviene come per il nostro giornalaio o salumiere che pensiamo di conoscere solo perché sono venti anni che facciamo acquisti da loro (ma poi non sappiamo nemmeno dove abitano o se sono sposati o no). Proviamo però ad essere meno superficiali ed a porci domande del tipo:

- che cosa sono i numeri?
- che cosa è un punto, una retta?
- i numeri, i punti le rette sono invenzioni dell'uomo, di un dio oppure esistono in natura?
- che cosa è la matematica?
- i risultati della matematica sono sicuri? e, se sono sicuri, perché lo sono?
- esiste l'infinito di cui spesso parla la matematica?
- perché la matematica, che non sembra avere a che fare con l'esperienza, è utile per le scienze empiriche e per le applicazioni?

(ho detto meno superficiali !!! chi avesse letto troppo rapidamente deve tornare indietro e riflettere su ciascuna domanda fino a quando non abbia raggiunto uno stato di confusione abbastanza avanzato). Ci accorgiamo allora che tali questioni sono molto più problematiche di quanto appare a prima vista. Ad esse persone diverse hanno dato e danno risposte diverse. Questo fatto si esprime dicendo che:

sono esistite ed esistono diverse *"filosofie"* della matematica.

Il mio punto di vista, che credo sia quello della maggioranza dei matematici che si interessano a queste cose, è che non abbia molto senso la pretesa di "fondare" la matematica una volta per tutte ma che invece sia interessante "viaggiare" tra le diverse intuizioni e pratiche dei matematici per vederne i collegamenti logici, le potenzialità ed i limiti. E' questo forse l'unico modo per tentare di capire qualche cosa di più sulla natura della matematica.

I due volumi non sono libri di storia della matematica anche se a volte si accenna a qualche riferimento storico come collante del discorso. In essi sono poi evidenti alcune lacune. Ad esempio lo spazio dedicato alla trattazione dell'intuizionismo non corrisponde certo alla importanza ed all'originalità di questa filosofia della matematica. Questa mancanza deriva dalla mia pigrizia e non dal misconoscimento di questo argomento. Inoltre il ruolo della logica nelle questioni dei fondamenti richiederebbe da solo lo spazio di un paio di volumi.

Insomma, per un discorso più approfondito e più completo si rimanda a libri più "solidi" di questi.

Salerno 2019

P.S. : Il mio indirizzo è ggerla@unisa.it e ricevo volentieri commenti, segnalazioni di errori o richieste di chiarimenti. L'indirizzo della mia homepage, dove può essere trovato un po' di materiale scientifico e didattico, è:
http://www.dipmat.unisa.it/people/gerla/www/

CAPITOLO 1

LA MATEMATICA PRESSO I GRECI

E' indegno del nome di uomo chi ignora il fatto che la diagonale di un quadrato è incommensurabile con il suo lato.
Platone (429-347 a.C.)

1. La scuola Pitagorica: tutto il mondo è numero (naturale)

Il primo organico tentativo di dare una fondazione alla matematica (ed all'intera conoscenza scientifica) fu probabilmente quello della scuola pitagorica il cui assunto di partenza era che:

alla base di tutto è il numero intero.

La scuola pitagorica era una setta mistico-religiosa che si sviluppò in Grecia e in Italia (Crotone) tra il 570 ed il 500 a.C. attorno ad un mitico personaggio chiamato Pitagora. Le idee di tale scuola sono di fondamentale importanza per la storia della cultura occidentale perché da esse inizierà quel processo che trasformerà la scienza pre-ellenica, che consisteva in una disarticolata raccolta di risultati dettati dall'esperienza, in una scienza razionale.[1] Dei pitagorici parla Aristotele al modo seguente, dove si deve tenere conto che allora per "numero" si intendeva "numero intero positivo".

> *Tra i primi filosofi, ..., furono i cosiddetti Pitagorici, i quali, applicatisi alle scienze matematiche, le fecero per i primi progredire; cresciuti poi nello studio di esse, vennero nell'opinione che i loro principi fossero i principi di tutti gli esseri... Pensarono che gli elementi dei numeri fossero gli elementi di tutte le*

[1]Non bisogna avere una immagine dei pitagorici come scienziati campioni di razionalismo. Il carattere mistico di questa scuola era fortissimo, siamo in presenza di una vera e propria setta religiosa (e politica) che credeva, tra le altre cose, che le anime dei morti si reincarnassero negli animali. Anche le "regole" di tale setta ci appaiono notevolmente bizzarre. Ad esempio ecco alcuni comandamenti:

- *non toccare un gallo bianco*
- *non addentare una pagnotta intera*
- *non guardare uno specchio accanto ad un lume.*

Ma queste che ci appaiono come stranezze non tolgono ai pitagorici il merito di costituire il punto di inizio della moderna cultura scientifica.

> *cose, e che l'universo intero fosse armonia e numero* (Aristotele, *Metafisica*).

Si deve tenere conto che in quel periodo era forte il desiderio di trovare i "principi ultimi" e che questi venivano cercati negli elementi naturali come l'aria, l'acqua o il fuoco. Forse però per capire meglio il pensiero dei Pitagorici conviene vedere cosa dice uno di loro, Filolao.

> *"Nessuna menzogna accolgono in sé la natura del numero e l'armonia: non è cosa loro la menzogna. La menzogna e l'invidia partecipano della natura dell'illimitato, dell'inintelligibile e dell'irrazionale. Nel numero non penetra menzogna, perché la menzogna è avversa e nemica della natura, così come la verità è connaturata e propria alla specie dei numeri . . . "*

> *"... Nulla sarebbe comprensibile, né le cose in sé né le loro relazioni, se non ci fossero il numero e la sua sostanza."*

> *"Tutte le cose che si conoscono hanno numero: senza il numero non sarebbe possibile pensare né conoscere alcunché."*

Un ruolo talmente centrale dato al numero potrebbe anche dipendere dalla scoperta di un forte collegamento tra rapporti numerici ed "armonia" in campo musicale. Infatti viene attribuita ai Pitagorici la scoperta di una scala armonica che viene detta, appunto, *scala pitagorica*. Consideriamo delle corde tese di varia lunghezza ed esaminiamo i suoni che vengono emessi pizzicandone due contemporaneamente. Ci si accorge che a volte si hanno effetti gradevoli ed a volte sgradevoli. E' possibile studiare quale sia il rapporto tra le lunghezze delle due corde ed il fenomeno della "gradevolezza" o, per essere più specifici, della "consonanza". Ora la prima scoperta che viene da fare è che se una corda è il doppio dell'altra si ha una fortissima consonanza. In questo caso noi diciamo che i due suoni differiscono per una ottava. Se indichiamo con A la lunghezza della prima corda e con B quella della seconda allora $B = (1/2)\cdot A$ o, se si vuole, $A : B = 2 : 1$. Un altro suono gradevole si ottiene facendo vibrare, insieme ad A una corda la cui lunghezza C sia i due terzi di A cioè $C = (2/3)\cdot A$. Ne segue che $A : C = 3 : 2$. Infine ci si accorge che un suono gradevole si ottiene dai suoni prodotti dalle corde C e B che risultano essere nel rapporto $C{:}B = 4{:}3$. Abbiamo quindi che le tre consonanze principali, (che prendono il nome di ottava quinta e quarta), corrispondono ai rapporti 2:1; 3:2 e 4:3. E' una sorprendente corrispondenza tra suoni e numeri che suggerisce fortemente l'idea per cui "tutto è numero".

Da quel *"tutte le cose che si conoscono hanno un numero"* scaturiva poi il convincimento circa la struttura granulare e discreta delle figure geometriche e, più in generale, del mondo fisico. Ciò comportava, ad esempio, una concezione del segmento come insieme finito di punti-unità, punti che venivano intesi come veri e propri corpi con una determinata grandezza. Infatti era solo in base a tale ipotesi che i numeri interi potevano essere lo strumento perfettamente adeguato alla descrizione della realtà, anzi, in un certo senso, venivano a coincidere con la realtà stessa. In tale modo la geometria non si considerava distinta dall'aritmetica e, in un certo senso, l'aritmetica assumeva una forma geometrica. Dei numeri infatti si dava una rappresentazione geometrica o, se si vuole, fisica, tramite una opportuna configurazione di punti-sassolino. Ad esempio si chiamavano *triangolari* i numeri che si potevano ottenere disponendo i punti in triangoli.

```
O      O       O          O
      O O     O O        O O
             O O O      O O O
                       O O O O
```

Si chiamavano invece *quadrati* i numeri corrispondenti a gruppi di sassolini disposti in quadrato

```
O     O O     O O O     O O O O
      O O     O O O     O O O O
              O O O     O O O O
                        O O O O
```

I numeri quadrati corrispondono ai numeri che sono quadrati perfetti. Ancora, si chiamavano *rettangolari* i numeri corrispondenti a gruppi di sassolini disposti in un rettangolo (che non si riduca ad una striscia di sassolini). In questo caso sono rettangolari tutti

```
O O     O O O     O O O     O O O O
O O     O O O     O O O     O O O O
                  O O O     O O O O
```

e soli i numeri che <u>non</u> sono numeri primi.

L' interpretazione dei numeri come particolari disposizioni di sassolini consente di sviluppare una interessante aritmetica. Ad esempio, è immediato che ogni triangolo si ottiene dal precedente aggiungendo un fila di sassolini. Pertanto se $t(n)$ è il numero dei sassolini dell'ennesimo triangolo abbiamo che la funzione t è definibile tramite le equazioni

$$t(1) = 1 \quad : \quad t(n) = t(n\text{-}1)+n.$$

Ne segue che sono triangolari tutti i numeri della serie 1, 3, 6, 10, ... di termine generale n, cioè tutti i numeri del tipo 1+2+3+...+n. Per quanto riguarda i numeri quadrati, è immediato vedere che ogni quadrato si ottiene dal precedente aggiungendo due lati (con un sassolino in comune). Ne segue che, se $q(n)$ è il numero dei sassolini dell'ennesimo quadrato, la funzione q si definisce tramite le equazioni

$$q(1) = 1 \quad : \quad q(n) = q(n\text{-}1)+2n\text{-}1.$$

Pertanto i quadrati perfetti si ottengono come elementi della serie 1, 1+3, 1+3+5, ..., 1+3+...+2n-1, cioè la serie di termine generale 2n-1. Si osservi che sia la funzione t che la funzione q sono state definite “per ricorsione”. In proposito si veda nel prossimo capitolo i paragrafi su induzione e ricorsione.

L'importanza della svolta impressa dalla scuola pitagorica non si limita alla sola matematica poiché la fede nella potenza regolarizzatrice del numero intero, il procedimento di astrazione, l'uso delle dimostrazioni nel procedere scientifico, rappresentano il nascere dell'aspetto fondamentale della cultura occidentale: il convincimento che il mondo sia comprensibile non attraverso l'ascesi mistica, la contemplazione, come viene ritenuto dalle culture orientali, ma attraverso l'attività raziocinante. Con Pitagora ha inizio un processo di idealizzazione e razionalizzazione di tutte le forme di conoscenza che dominerà perfino la nostra cultura religiosa. Afferma ad esempio Bertrand Russell in *"Storia della filosofia occidentale"* :

“La religione razionalistica, al contrario di quella apocalittica, è stata da Pitagora in poi (ed in particolare da Platone in poi) completamente dominata dalla matematica e dal metodo matematico. La combinazione di matematica e di teologia, che cominciò con Pitagora, caratterizzò la filosofia religiosa in Grecia, nel Medioevo e nell'era moderna fino a Kant. L'orfismo precedente a Pitagora era analogo alle misteriose religioni asiatiche. Ma, in Platone, Sant'Agostino, Tommaso d'Aquino,

Cartesio, Spinosa e Leibniz, vi è un intimo intrecciarsi di religione e di ragionamento, di aspirazione morale e di ammirazione logica per ciò che è eterno, il quale viene da Pitagora e distingue la teologia intellettualizzata dell'Europa dal più diretto misticismo asiatico."

2. I numeri non sono sufficienti: grandezze incommensurabili
Abbiamo visto che per la scuola pitagorica ogni figura è costituita da un certo numero di sassolini-punto disposti opportunamente. Se si accetta questo punto di vista ogni figura geometrica, in un certo senso, può essere misurata da un numero naturale. Infatti basta contare il numero dei sassolini-punto che ne fanno parte. Purtroppo vedremo che questo non è sempre possibile. Per prima cosa vediamo che cosa significa misurare un segmento.

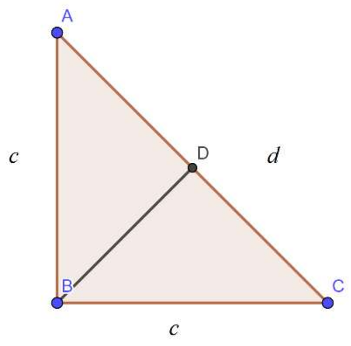
In altre parole, dovendo misurare un segmento a e l'unità di misura è un segmento u, allora dico che la misura di a è n se "riportando" *n volte u* lungo il segmento a riesco a raggiungere esattamente la fine del segmento a.

Definizione 2.1. Diciamo che un segmento a è un *n-multiplo* di un segmento u se a si può dividere in n segmenti *congrui* a b. Se assumiamo u come segmento unitario, allora diciamo che la *misura di a* rispetto all'unità di misura u è n.

In altre parole, dovendo misurare un segmento a e l'unità di misura è un segmento u, allora dico che la misura di a è n se "riportanto" u lungo *a n volte* riesco a raggiungere esattamente la fine del segmento a.

Definizione 2.2. Diciamo che i segmenti $a_1,\ldots,a_n$ sono *commensurabili* se esiste un segmento u tale che ogni a_i ha una misura intera rispetto all'unità di misura u. Altrimenti diciamo che sono *incommensurabili.*

Teorema 2.3. Dato un triangolo rettangolo isoscele, qualunque sia l'unita di misura scelta, cateto e diagonale non possono avere entrambi misura intera.



Dim. Fissata un'unità di misura *u*, supponiamo per assurdo che esistano triangoli rettangoli isosceli con cateto ed ipotenusa di misura intere. Sia *d* la lunghezza più piccola tra le lunghezze di tali triangoli ed indichiamo con ABC un triangolo la cui ipotenusa misuri *d* mentre supponiamo che i suoi cateti misurino *c*. Indichiamo con *D* il punto medio dell'ipotenusa AC. E' evidente che l'area del triangolo ABC è il doppio dell'area del triangolo BDC e quindi che $c^2/2 = 2\cdot(d^2/2)$. Da tale equazione segue che $d^2 = 2\cdot c^2$ e quindi che *d* è un numero pari. Poiché BD è uguale a DC, questo comporta che il triangolo BDC ha lati interi. Questo è assurdo in quanto BCD ha una ipotenusa *c* minore di quella *d* di ABC.

Poiché cateto e diagonale di un triangolo rettangolo isoscele possono essere anche visti come lato e diagonale di un quadrato, possiamo anche enunciare il teorema ora dimostrato al modo seguente.

Teorema 2.4. Il lato e la diagonale di un quadrato sono incommensurabili.

Possiamo anche riformulare quanto detto fino ad ora come un tipo di paradosso.

Paradosso 2.5. (Paradosso dell'esistenza di una figura con lati incommensurabili)[2] Dato un quadrato, per quanto piccola si sia

[2] Letteralmente "paradosso" significa "contro l'opinione comune" e tale espressione non andrebbe confusa con "antinomia" o "contraddizione" che invece significano che si è riusciti a dimostrare una affermazione ed anche la negata di tale affermazione. Tuttavia spesso si confondono le due cose anche perché, se si esprime l'opinione comune come assioma, allora il paradosso diviene una contraddizione. Anche l'espressione "opinione comune" può significare cose diverse e dipende molto dall'epoca o dal gruppo sociale cui si riferisce. Nel nostro caso il paradosso deve essere inteso come contro l'opinione comune ai tempi della

scelta l'unità di misura (centimetro, millimetro, decimillimetro ...) se si riesce a misurare con precisione il lato non è possibile misurare con precisione la diagonale.

Abbiamo visto quindi che il tentativo di fondare la geometria a partire dai numeri interi non regge. Si potrebbe pensare che estendendo la nozione di misura in modo da accettare i numeri razionali tale tentativo possa avere successo.

Definizione 2.6. Diremo che *u'* è la *m-sima parte di u* se *u* è un *m*-multiplo di *u'*. Diremo che la misura di un segmento *a* rispetto ad *u* è il numero razionale *p/q* se *u* è un *p*-multiplo di segmenti che sono congrui all'*q*-sima parte di *u*.

Teorema 2.7. Dato un quadrato, *ABCD* qualunque sia l'unità di misura scelta *u*, lato e diagonale non possono avere entrambi misura razionale.

Dim. Assumiamo per assurdo che lato e diagonale si possano rappresentare con due razionali e rappresentiamo questi due razionali con lo stesso denominatore e quindi con le frazioni n/t ed m/t. Ingrandiamo allora *ABCD* di *t*. Otteniamo allora un *A'B'C'D'* con lato e diagonale di lunghezza *n* ed *m*, rispettivamente in contraddizione col Teorema 2.4.

3. Il teorema di Pitagora

Lo sapevate? Il quadrato costruito sull'ipotenusa è il doppio di quello sui cateti . . . ma la qualità è scadente e dopo un anno lo butti! È così! È capitato a mia sorella! Fidatevi!
-Vulvia (Corrado Guzzanti, Il caso Scrafoglia, 2002).

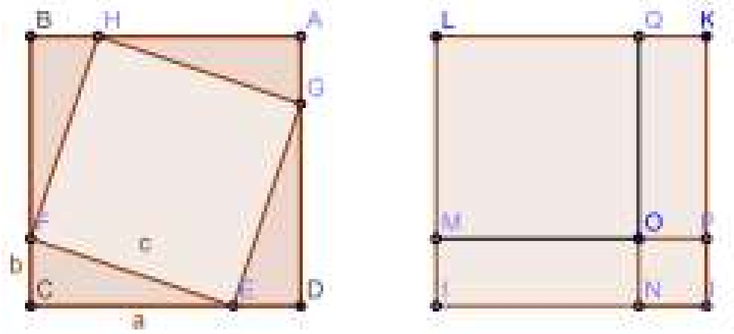
Dimostrazioni di esistenza di figure con lati non commensurabili sono rese più facili se si utilizza il teorema di Pitagora.

---

scuola Pitagorica oppure contro l'opinione comune della comunità delle persone non acculturate..

Teorema 3.1. (Teorema di Pitagora) Dato un triangolo rettangolo, se si considera l'unione dei due quadrati costruiti sui cateti otteniamo una figura che ha la stessa estensione del quadrato costruito sull'ipotenusa.[3]

Dim.[4] Indichiamo con T il triangolo rettangolo, supponiamo che i cateti misurino a e b e l'ipotenusa c. Costruiamo un quadrato Q con lati uguali ad $a+b$.



Detti A, B , C, D i vertici di Q, tracciamo sui lati i quattro punti H, F, E, G in modo che $CE = DC = AH = BF = a$ mentre $ED = GA = HB = FC = b$. In tale modo si individuano quattro triangoli rettangoli uguali (avendo cateti uguali per costruzione). Inoltre tali punti individuano un quadrilatero $FEGH$. Tale quadrilatero

[3] Attualmente tale teorema si enuncia dicendo che: "la somma dei quadrati delle misure dei cateti è uguale al quadrato della misura dell'ipotenusa". Tuttavia si deve tenere presente che per i greci non esisteva una "misura" di una figura geometrica in quanto misurare significa assegnare un numero reale a qualche cosa e nella matematica greca non esisteva la nozione di numero reale. Come vedremo, essi avevano invece la nozione di uguaglianza dell'estensione di due figure, quella di equiscomponibilità e quella di proporzionalità di grandezze omogenee con cui riuscivano ad esprimere molti teoremi sulle misure.

[4] In realtà quella che segue non è una dimostrazione nel senso moderno e rigoroso del termine e non lo può essere visto che non abbiamo ancora elencato gli assiomi della geometria di Euclide. Piuttosto si prova una proprietà non molto evidente "il teorema di Pitagora" a partire da altre proprietà che ci appaiono più evidenti. Ad esempio il primo passo della dimostrazione consiste nella costruzione di un quadrato di un dato lato. E' possibile sempre costruire un tale quadrato? La questione non è tanto semplice ed ha a che fare con l'assioma delle parallele. Nel corso di questo capitolo considereremo spesso "dimostrazioni" di questo tipo.

ha lati uguali in quanto coincidenti con le ipotenuse dei triangoli. Inoltre gli angoli sono retti. Ad esempio, l'angolo in *E* è retto in quanto è uguale ad un angolo piatto meno la somma dei due angoli non retti di *T*. D'altra parte la somma dei due angoli non retti di un triangolo rettangolo è un angolo retto. Pertanto *FEGH* è un quadrato, precisamente il quadrato costruito sull'ipotenusa..

Osserviamo ora che l'area di *Q* può essere calcolata come il quadrato del lato cioè come $(a+b)^2$ e quindi, per la formula del quadrato di un binomio, come a^2+b^2+2ab. D'altra parte l'area di *Q* è anche uguale a quella del quadrato piccolo *FEGH* più quella dei quattro triangoli, cioè è uguale a $c^2+(4ab)/2$. Dall'uguaglianza

$$a^2+b^2+2ab = c^2+2ab$$

si ricava che $a^2+b^2=c^2$.

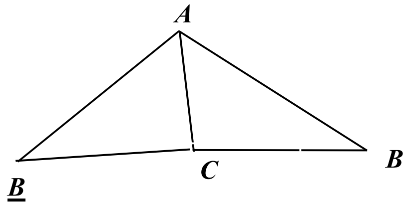
Avendo utilizzato la formula del quadrato del binomio una tale dimostrazione è di carattere algebrico. Tuttavia tale formula ammette una semplice dimostrazione geometrica che è completamente illustrata dalla figura tracciata sopra a destra in cui si mostra come l'area del quadrato si possa calcolare in due modi. Uno è considerare il quadrato del lato $a+b$, l'altro è effettuare la somma dei quadrati di estensione a^2 e b^2 più due rettangoli di estensione ab. □

E' naturale anche porsi il seguente problema:
"un triangolo che verifica il teorema di Pitagora è necessariamente rettangolo ?
Ad esempio, possiamo avere tre asticelle di lunghezza 3, 4, 5 e chiederci se il triangolo costruito con tali asticelle sia rettangolo. Il seguente teorema mostra che la risposta è positiva.

Teorema 3.2. (Teorema inverso di Pitagora) Ogni triangolo i cui lati verificano il teorema di Pitagora è rettangolo. [5]

[5] Questo teorema è utilizzato per costruire un angolo retto con una semplice corda. Basta individuare sulla corda quattro punti consecutivi *A*, *B*, *C*, *A'* in modo che i segmenti *AB*, *BC* e *CA'* siano di lunghezza 5, 4, 3, rispettivamente. Poi, congiungendo *A* con *A'* si costruisce il triangolo *ABC*. Tenendo conto del fatto che $3^2+4^2 = 5^2$ siamo sicuri che il triangolo è rettangolo in *C*.

Dim. Sia ABC un triangolo tale che $AB^2 = AC^2+BC^2$ e costruiamo un segmento $\underline{B}C$ perpendicolare ad AC e di lunghezza uguale a CB. Allora i due triangoli $AC\underline{B}$ e ACB hanno il lato AC in comune ed i lati $\underline{B}C$ e CB uguali per costruzione. Inoltre, essendo $AC\underline{B}$ rettangolo in C, $A\underline{B}^2 = AC^2+\underline{B}C^2 = AC^2+CB^2 = AB^2$.



Pertanto i due triangoli, avendo i tre lati uguali, sono uguali. Da ciò segue che l'angolo ACB è uguale all'angolo retto $AC\underline{B}$ ed è quindi retto. □

4. Ancora grandezze incommensurabili: la prima crisi

Il teorema di Pitagora è un valido strumento per trovare infiniti esempi di grandezze incommensurabili. Per poterlo fare si serviremo dei due seguenti lemmi.

Lemma 4.1. Se c è un quadrato perfetto, allora ogni numero primo presente nella fattorizzazione di c ha un esponente pari.

Dim. . Supponiamo che $c = n^2$ e che n abbia la scomposizione $n = p_1^{n(1)},\ldots,p_s^{n(s)}$ con $p_1,\ldots,p_s$ primi allora $c = n^2 = p_1^{2\cdot n(1)}\cdot\ldots\cdot p_s^{2\cdot n(s)}$.

Supponiamo ad esempio che $c = 12^2 = 144$ allora essendo $12 = 2^2\cdot 3$, la fattorizzazione di 144 sarà $12^2 = 2^4\cdot 3^2$.

Lemma 4.2. Sia p un numero primo, allora un quadrato perfetto non può essere p volte un altro quadrato perfetto, cioè l'equazione

$$p\cdot n^2 = m^2$$

non ammette soluzioni intere.

Dim. Supponiamo per assurdo che esistano due interi n ed m tali che $p\cdot n^2 = m^2$, e poniamo $c = p\cdot n^2 = m^2$, allora

- m^2 e quindi m è divisibile per p,
- dall'equazione $c = m^2$ si deduce che il fattore p compare in c un numero pari di volte

- dall'equazione $c = p{\cdot}n^2$, tenendo conto del fatto che in n^2 il fattore p compare con esponente pari, si deduce che il fattore p è presente in c un numero dispari di volte.

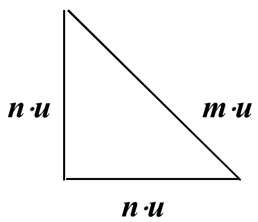
Questa conclusione è un assurdo e quindi il lemma è provato. □

Da notare che l'equazione $p{\cdot}n^2 = m^2$ equivale a dire che $p = m^2/n^2 = (m/n)$, cioè equivale a dire che il numero razionale m/n è la radice di p. Quindi il lemma 4.2 equivale alla seguente proposizione

Proposizione 4.3. La radice di un numero primo non è un razionale.

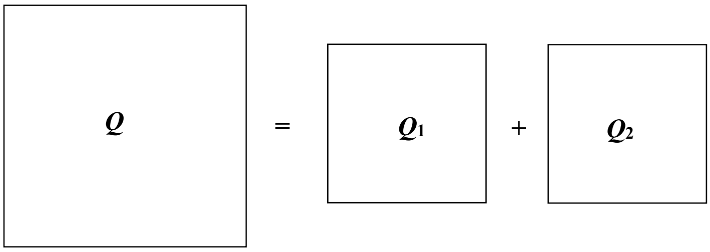
Teorema 4.4. (**Paradosso dell'esistenza di grandezze non commensurabili**) Dato un triangolo rettangolo isoscele, qualunque sia l'unita di misura scelta, cateto e diagonale non possono avere entrambi misura intera.

Dim. Supponiamo per assurdo che esista un segmento u (unità di misura) tale che sia il lato che la diagonale siano multipli, secondo gli interi n ed m, di u. In altri termini, supponiamo che il lato e la diagonale siano misurati tramite due numeri interi n ed m. Allora i quadrati costruiti sui cateti sono costituiti da n^2 quadratini unitari mentre il quadrato costruito sull'ipotenusa è costituito da m^2 quadratini unitari. Ne segue che per il teorema di Pitagora $n^2 + n^2 = m^2$ e quindi $2n^2 = m^2$. Ma ciò, come abbiamo visto nel lemma precedente, è assurdo. □



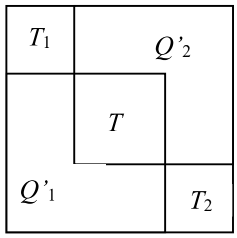
L'incommensurabilità tra lato e diagonale del quadrato può anche essere riformulata al modo seguente. Fissata un' unità di misura chiamiamo *quadrato intero* un quadrato il cui lato abbia misura intera.

Teorema 4.5. (Conway). Non esiste un quadrato intero che sia somma di due quadrati interi uguali. In altri termini non è possibile una configurazione del tipo



con $Q_1 = Q_2$ e Q_1, Q_2 e Q quadrati interi.

Dim. Tale teorema non andrebbe dimostrato essendo solo una riformulazione del teorema 4.2. Tuttavia esiste la seguente bella dimostrazione dovuta al grande matematico Conway. Indichiamo con P la proprietà per cui Q è un quadrato intero somma di due quadrati interi Q_1 e Q_2 uguali tra loro. Allora possiamo considerare il quadrato Q di lato minimo verificante P. Spostiamo ora Q_1 e Q_2 all'interno di Q in modo da ottenere



- due quadrati sovrapposti Q'_1 e Q'_2,
- due quadratini T_1 e T_2
- un quadrato intersezione $T = Q'_1 \cap Q'_2$.

Essendo le lunghezze dei lati di Q, Q_1, Q_2 intere, i lati di tali figure hanno lunghezze intere. Indicando con $\mu(X)$ l'area di una figura X, risulta che

$$\begin{aligned}\mu(Q) &= \mu(Q'_1 \cup Q'_2 \cup T_1 \cup T_2)\\ &= \mu(Q'_1 \cup Q'_2) + \mu(T_1) + \mu(T_2)\\ &= \mu(Q'_1) + \mu(Q'_2) - \mu(T) + \mu(T_1) + \mu(T_2)\\ &= \mu(Q_1) + \mu(Q_2) - \mu(T) + \mu(T_1) + \mu(T_2).\end{aligned}$$

Da ciò si ricava che $\mu(T) = \mu(T_1) + \mu(T_2)$ e quindi che T verifica P. Essendo T più piccolo di Q, ciò contraddice l'ipotesi per cui Q è il quadrato più piccolo verificante P. □

Un altro modo di provare l'incommensurabilità è il seguente dove chiamiamo *triangolo impossibile* un triangolo rettangolo isoscele (con cateti uguali) con i tre lati interi.

Teorema 4.6. Dato un triangolo impossibile possiamo costruirne un altro con area dimezzata.

Dim. Sia DBC un triangolo impossibile, chiamiamo con d la lunghezza dell'ipotenusa e con c quella dei cateti. Allora poiché per il teorema di Pitagora $d^2 = 2 \cdot c^2$, d è pari. Ne segue che se D' è il punto di mezzo di DB, la lunghezza di D'B è un intero. Ne segue che *D'BC* è un triangolo impossibile.

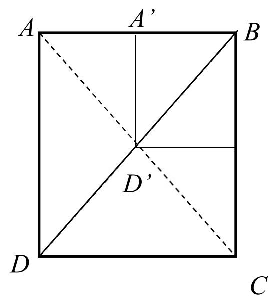
Possiamo riformulare lo stesso teorema al modo seguente dove per *quadrato impossibile* si intende quadrato con lato intero e diagonale intera.

Teorema 4.7. Ogni quadrato impossibile si può tagliare in quattro quadrati impossibili.

Poiché non è possibile costruire una successione di interi strettamente decrescente, vale il seguente corollario.

Corollario 4.8. Non esistono triangoli o quadrati impossibili.

Abbiamo visto che i numeri interi non costituiscono uno strumento sufficiente per effettuare misure geometriche. La cosa non è poi tanto sorprendente visto che, come vedremo, ogni segmento può essere dimezzato. Quindi se un segmento ha una lunghezza dispari la sua metà è un numero razionale. Si potrebbe allora pensare che le cose migliorano se si coinvolgono i numeri razionali (che d'altra parte non venivano accettati dagli antichi greci).



La crisi. Abbiamo visto che, detto in termini attuali, perfino la misurazione di grandezze legate alla figura geometrica più elementare, il quadrato, comporta necessariamente il coinvolgimento dei numeri irrazionali. Ora, se si tiene conto del fatto che per gli antichi greci i soli numeri esistenti erano gli interi positivi, ciò significava per la cultura del tempo che:

> vi sono più cose in geometria (e quindi nel mondo fisico) di quanto i numeri siano capaci di esprimere !

D'altra parte, poiché la razionalità veniva identificata con il numero (cioè la possibilità di esprimere tramite numeri e rapporti di numeri i fenomeni del mondo), una tale conclusione comportava la messa in discussione della stessa possibilità, da parte dell'uomo, di pervenire alla conoscenza. Un altro fatto notevole è che:

> alla luce della scoperta di grandezze incommensurabili, come il lato e la diagonale del quadrato, la stessa concezione dei segmenti ed in generale di ogni figura geometrica come somma finita di punti-atomo, che era propria della scuola pitagorica, non poteva più reggere. [6]

Se infatti i segmenti fossero costituiti da un numero finito di particelle indivisibili, tutte di lunghezza u, allora il lato e la diagonale del quadrato avrebbero lunghezza multiplo di u e sarebbero quindi commensurabili. D'altra parte una tale concezione coincide proprio con quella della fisica moderna che sembra pertanto essere in contrasto con la geometria, o, per meglio dire, con il teorema di Pitagora.

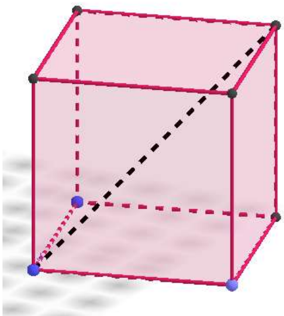
Concludiamo osservando che:

---

[6] Come fosse sconvolgente per i greci una tale scoperta possiamo vederlo attraverso questo passo attribuito al filosofo Proclo Diadoco.

> *"E' fama che colui il quale per primo rese di pubblico dominio la teoria degli irrazionali sia perito in un naufragio, e ciò perché l'inesprimibile e l'inimmaginabile sarebbero dovuti rimanere sempre celati. Perciò il colpevole, che fortuitamente toccò e rivelò questo aspetto delle cose viventi, fu trasportato al suo luogo di origine e là viene in perpetuo flagellato dalle onde."*

se con la scuola pitagorica abbiamo il primo tentativo di fondazione generale della matematica, con la scoperta delle grandezze incommensurabili siamo in presenza della prima "crisi dei fondamenti" della matematica.



Problema: Consideriamo un cubo e chiamiamo *diagonale* il segmento che congiunge due vertici opposti. Verificare che la diagonale non è commensurabile con lo spigolo.

Radici quadrate e razionali. E' evidente che la questione delle grandezze incommensurabili è legata alle radici quadrate di numeri interi. Ad esempio chi legge può provare a risolvere i seguenti problemi utilizzando le stesse tecniche utilizzate nei paragrafi precedenti per l'incommensurabilità.

Problema. Dimostrare (per assurdo) che la radice di due non è razionale. Si suggerisce di seguire la linea di dimostrazione del teorema in cui si prova l' incommensurabilità della diagonale con il lato di un quadrato.

Problema. Dimostrare che, dato un numero primo p, la radice di p non è razionale.

Problema. Sia n il prodotto di numeri primi diversi tra loro. Dimostrare che la radice di n non è razionale.

5. Dimostrare e dimostrare per assurdo

Nei paragrafi precedenti abbiamo visto alcuni tra i più antichi esempi di "dimostrazione". Tali dimostrazioni sono di natura molto diversa. Nel teorema di Pitagora si prova qualcosa in positivo, precisamente che vale una certa equazione e la dimostrazione consiste nel mettere in evidenza che valgono una serie di uguaglianze. La dimostrazione della incommensurabilità del lato e della diagonale di un quadrato è invece diversa e costituisce uno dei primi esempi di *"dimostrazione per assurdo"*. In questo caso non si prova qualcosa ma si prova che non può valere qualcosa. Stante l'importanza di tale tipo di dimostrazione, esaminiamo da vicino la sua struttura.

Ricordiamo che chiamiamo *contraddizione* l'affermazione di un fatto e la contemporanea negazione di tale fatto. Indicando con B una asserzione, con $\neg B$ la sua negata, con $\wedge$ la congiunzione logica "e", allora una contraddizione ha una forma del tipo $B\wedge(\neg B)$ che si legge "B e non B". Naturalmente una contraddizione non può essere un'asserzione vera. Inoltre, poiché da asserzioni vere si deducono ancora asserzioni vere:

se *da una asserzione A segue una contraddizione,*
allora *A non può essere vera,*
pertanto *: ¬A è vera.*

In definitiva la struttura di una dimostrazione per assurdo di una proposizione $\neg A$ è questa:

1. si suppone A
2. da tale ipotesi si ricava una contraddizione $B\wedge\neg B$,
3. si conclude che, non potendo valere A, vale $\neg A$.

Ad esempio, nel caso della dimostrazione di incommensurabilità del lato e della diagonale del quadrato,
- si suppone la commensurabilità,
- da tale ipotesi si ricava sia l'affermazione B = " in c il fattore 2 è presente un numero pari di volte" sia l'affermazione $\neg B$ = " in c il fattore 2 è presente un numero dispari di volte".
- si conclude che non sussiste la commensurabilità.

Naturalmente le dimostrazioni per assurdo possono essere utilizzate anche per provare qualche cosa in positivo. Possiamo infatti affermare il seguente altro principio:

se *da una asserzione ¬A segue una contraddizione,*
allora *¬A non può essere vera*
pertanto*: A è vera.*

Pertanto, dovendo dimostrare A,
1. si suppone $\neg A$
2. da tale ipotesi si ricava una contraddizione $B\wedge\neg B$,
3. si conclude che, non potendo valere $\neg A$, vale A.

Un'altra bella dimostrazione per assurdo è la dimostrazione per cui i numeri primi sono infiniti.

Teorema 5.1. L'insieme P dei numeri primi è infinito.

Dim. Detto A l'enunciato *"P è infinito"* neghiamolo, cioè supponiamo per assurdo che l'insieme P dei numeri primi sia finito. Posto $P = \{p_1,... p_n\}$, sia $q = p_1\cdot...\cdot p_n +1$: vogliamo dimostrare che q è primo. Infatti se per assurdo q non fosse primo ammetterebbe un divisore primo $p \neq 1$. Poiché abbiamo supposto che $p_1,...,p_n$ sono tutti i possibili numeri primi, p deve coincidere con un opportuno p_i. Allora p_i, dividendo sia q che $p_1\cdot...\cdot p_n$ dovrà dividere anche q-$p_1\cdot...\cdot p_n = 1$, cosa questa assurda. L'assurdo a cui siamo pervenuti prova che q è primo. D'altra parte q, essendo maggiore dei numeri $p_1,..., p_n$, non appartiene a $\{p_1,...,p_n\}$, in contrasto con l'ipotesi che $\{p_1,..., p_n\}$ è l'insieme di tutti i numeri primi. □

Per provare questo teorema abbiamo dovuto utilizzare due volte il metodo di dimostrazione per assurdo. Possiamo eliminare una di queste utilizzazioni e riformulare il teorema in modo più "costruttivo".

Teorema 5.2. Dato un insieme finito $\{p_1,...,p_n\}$ di numeri primi esiste un numero primo p che non appartiene a $\{p_1,...,p_n\}$. Ne segue che l' insieme dei numeri primi è infinito.

Dim. Consideriamo il numero $q = p_1\cdot...\cdot p_n +1$ e sia p un divisore di q diverso da 1 (che potrebbe coincidere con q se q fosse primo). Se per assurdo $p\in\{p_1,...,p_n\}$ allora p, dividendo sia q che $p_1\cdot...\cdot p_n$, dividerebbe anche q-$p_1\cdot...\cdot p_n = 1$ e ciò è assurdo. □

Ad esempio, dato l'insieme {2, 3, 7} di numeri primi, il numero $2\cdot3\cdot7+1 = 43$ è primo. Dato l'insieme {3,7, 11} di numeri primi, il numero $3\cdot7\cdot11 +1 = 232$ non è primo ma è divisibile per il numero primo 2 che è diverso da 3, 7, 11.

Esempio. Quando si studiano le equazioni si dimostrano spesso "teoremi" per assurdo. Ad esempio proviamo il seguente "teorema":

A = "L'equazione $x^2+2\cdot(1+x^2) = 3x^2-2$ non ammette soluzioni."
Se neghiamo A dobbiamo accettare l'esistenza di un elemento r tale che

$r^2+2\cdot(1+r^2) = 3r^2-2$

da ciò seguirebbe che

$r^2+2+2\cdot r^2 = 3r^2-2$ e quindi

$3\cdot r^2+2 = 3r^2-2$ e quindi, semplificando $3\cdot r^2$

$2+2 = 0$ cioè

$4 = 0$.

L'assurdo a cui siamo pervenuti ci assicura che vale A.[7]

Problema. Consideriamo il numero n = 100!+1, dire quali delle seguenti affermazioni è esatta:
1. n è divisibile per 26, 2. n è divisibile per 65,
3. n è divisibile per 81, 4. Nessuna delle risposte precedenti.

Problema. Quale è il più piccolo numero divisibile per i primi cento numeri?

6. Il continuo geometrico per evitare l'infinito attuale

Il fatto che i numeri interi si rivelassero uno strumento inadeguato a definire le grandezze geometriche poteva, da un punto di vista tecnico, essere risolto (almeno) in due modi diversi.

1. Si poteva ampliare il concetto di numero.
2. Si poteva decidere che la geometria non è riducibile all'algebra, cioè alla nozione di numero.

I Greci seguirono la seconda via. Il primo punto di vista sarà invece assunto, come vedremo nel seguito, dalla matematica moderna con la definizione dei numeri reali e la successiva loro utilizzazione per la costruzione del continuo geometrico (la famosa geometria analitica).[8] Poiché i reali si definiscono a partire dai

---

[7] In realtà non siamo pervenuti ad un assurdo poiché l'equazione 4 = 0 è da considerare assurda solo se siamo nell'ambito di teorie in cui sia dimostrabile $\neg(4=0)$ come avviene nella teoria dei numeri reali. Se invece siamo, ad esempio, nell'anello degli interi modulo 4 la conclusione non è affatto assurda.

[8] I Greci accettavano solo i numeri naturali. Gli stessi numeri razionali erano considerati a volte come operatori, a volte come relazioni tra grandezze. Ad esempio quello che per noi è il numero 3/4, per i greci non era un ente matematico in qualche modo esistente ma solo un modo abbreviato per dire "prendi una grandezza, moltiplicala per 3 e dividila

razionali, e questi a partire dagli interi, l'attuale punto di vista sembra il naturale sviluppo di quello della scuola pitagorica. D'altra parte i Greci non potevano definire i reali perché non è possibile definirli a partire dagli interi senza coinvolgere la nozione di infinito attuale. Basta osservare che, come faremo in seguito, un numero reale si definisce come un insieme attualmente infinito di razionali (si veda il metodo delle sezioni) oppure come una successione attualmente infinita di cifre decimali.

Ma gli antichi greci rifiutavano l'infinito attuale

> *“. . . ché il numero è infinito in potenza, ma non in atto ... questo nostro discorso non intende sopprimere per nulla le ricerche dei matematici per il fatto che esso esclude che l'infinito per accrescimento sia tale da poter essere percorso in atto. In realtà essi stessi (i matematici), allo stato presente, non sentono il bisogno dell'infinito (e in realtà non se ne servono) ma soltanto di una quantità grande quanto essi vogliono, ma pur sempre finita ... “* (Aristotele).

Ancora, Aristotele (Fisica III) afferma che l'infinito è tale

> *“... che si può prendere sempre qualcosa di nuovo (in esso), e ciò che si prende è sempre finito ma sempre diverso. Sicché non bisogna prendere l'infinito come un singolo essere, per esempio un uomo o una cosa, ma nel senso in cui si parla di una giornata o di una lotta, il cui modo d'essere non è una sostanza ma un processo e che, se pure è finito, è incessantemente diverso.”*

Esistono mille esempi in cui si manifesta questo rifiuto dell’infinito da parte dei Greci. Ad esempio, il teorema da noi dimostrato

per 4". Altre volte 3/4 stava ad indicare un certa relazione tra due grandezze, infatti aveva senso scrivere, date due grandezze a e b, che a e b sono nella proporzione di 3 a 4, in breve $a : b = 3 : 4$ (si veda la teoria delle proporzioni). Un tale modo di vedere i razionali comportava poi difficoltà a definire le usuali operazioni di addizione e moltiplicazione. Infatti sembrava difficile giustificare l'addizione o la moltiplicazione di due operazioni o di due relazioni. A maggior ragione per i greci era inconcepibile una teoria degli irrazionali come quella attuale. Essi avevano però una tecnica, la teoria delle grandezze omogenee, che, come vedremo, permetteva ugualmente di esprimere quei concetti che, per i matematici moderni, coinvolgono gli irrazionali.

circa l'esistenza di infiniti numeri primi in realtà veniva enunciato dai Greci al modo seguente in cui non viene coinvolta la nozione di insieme infinito.

Per ogni primo p esiste un primo q maggiore di p.

Inoltre non si concepiva la retta come qualche cosa di illimitato come facciamo oggi. Si faceva riferimento solo ai segmenti ammettendo tuttavia che un segmento si possa prolungare a piacere.

Il motivo per cui i Greci avessero tanta repulsione per l'infinito attuale è probabilmente di natura strettamente filosofica ed è ben illustrato dal passo dei pitagorici che abbiamo citato in cui si afferma che la menzogna e l'invidia partecipano della natura dell'illimitato. Comunque una certa influenza deve avere pure avuto il fatto che esistevano paradossi legati all'infinito attuale che dimostravano le difficoltà logiche di tale concetto. Forse il paradosso più famoso è quello, di "Achille e la tartaruga" dovuto a Zenone.

Paradosso di Achille e la tartaruga. In questo paradosso si racconta di una sfida di una tartaruga (simbolo della lentezza) ad Achille (noto per la sua velocità) in una corsa. In tale sfida la tartaruga dichiara che purché gli siano dati dieci metri di vantaggio, non si sarebbe fatta raggiungere da Achille. Achille accetta la sfida, partono insieme ed Achille percorre quei dieci metri di vantaggio. Nel frattempo tuttavia la tartaruga percorre un metro; Achille non si scoraggia ed allora percorre quel metro, ma nel frattempo la tartaruga percorre un decimetro; Achille percorre quel decimetro, ma nel frattempo la tartaruga percorre un centimetro . . . e così via all'infinito. In questo modo Achille può correre per sempre senza raggiungere mai la tartaruga.[9]

---

[9]Attualmente tale paradosso viene "risolto" coinvolgendo la nozione di serie convergente. Infatti gli infiniti intervalli impiegati ogni volta da Achille per raggiungere la tartaruga diventano non solo sempre più piccoli ma il limite della loro somma *converge*. Di questa possibile soluzione (che comunque non è universalmente accettata ed anche a me non sembra centrare il problema) era convinto anche Cartesio come mostra il seguente passo in cui, riferendosi al paradosso afferma che:

"non è difficile a risolversi, quando si consideri che alla decima parte di una quantità viene aggiunta la decima di questa decima, e cioè una centesima; e poi ancora la decima di quest'ultima, ossia una millesima della prima; e così di seguito all'infinito, tutte queste decime

7. Punti linee e Platonismo

Come abbiamo già detto, dalla scoperta delle grandezze incommensurabili in poi la geometria assume un ruolo centrale nella conoscenza scientifica e filosofica dell'antica Grecia. Essa ha uno sviluppo che appare enorme se lo si confronta con le altre branche della conoscenza. Per rendersene conto basta pensare che la geometria che si impara a scuola è solo una piccolissima parte della geometria scoperta di greci (se si esclude la geometria analitica che è una scoperta relativamente recente). Al contrario, gli argomenti di fisica, chimica, biologia che fanno parte dei programmi scolastici sono enormemente superiori per quantità e qualità a quelli che la persona più istruita dell'antica Grecia poteva possedere.

Esaminiamo gli aspetti più rilevanti della geometria dei Greci e partiamo da quello sicuramente più importante: l'idealizzazione degli enti geometrici.

<u>**Primo processo di idealizzazione.**</u> Un primo processo di idealizzazione consiste nel dare carattere di "sostanza" a quelle che prima erano considerate proprietà della materia. Se nella matematica pre-ellenica "essere quadrato" era un attributo di alcuni oggetti materiali, non diverso da "essere rosso", "essere pesante", nel seguito si perverrà ad un nuovo ente "il quadrato" che a tutti gli effetti verrà trattato come una sostanza individuale. Questo significa che un quadrato diviene un oggetto di cui è possibile descrivere le proprietà allo stesso modo di come viene fatto per tutte le cose esistenti in natura. Dal punto di vista grammaticale,

---

prese insieme, benché siano supposte realmente infinite, non compongono tuttavia che una quantità finita. Ché se taluno dice che una tartaruga, la quale ha dieci leghe di precedenza rispetto a un cavallo dieci volte più veloce di lei, non potrà mai essere superata da questo, perché mentre il cavallo compie le dieci leghe la tartaruga ne percorre una e, mentre il cavallo supera questa lega, la tartaruga procede ancora di un decimo di lega e così all'infinito, bisogna rispondere che veramente il cavallo non la sopravanzerà finché esso farà quella lega, quel decimo, quel centesimo, quel millesimo ecc. di lega; ma che non ne segue che non la supererà mai, perché quel decimo, centesimo, millesimo ecc. non fanno che un nono di lega, in capo al quale il cavallo comincerà a sopravanzarla" (*Lettres de Mr Descartes*, Paris, 1657, N. 118).

questo fenomeno si manifesterà nella trasformazione del ruolo di parole come "quadrato", "punto", "segmento" le quali da attributi divengono soggetti. Così a frasi del tipo "quel tavolo è un quadrato", che pongono in relazione un ente materiale (quel tavolo) con una sua possibile proprietà (essere quadrato) che appartengono, per così dire, alla fisica, si vengono sostituendo espressioni del tipo "il quadrato ha le diagonali uguali". Tali espressioni pongono in relazione enti e proprietà ideali e la loro validità, non potendo dipendere dalla esperienza del mondo esterno, può essere stabilita solo all'interno di una organizzazione razionale delle conoscenze.

Secondo processo di idealizzazione. Un secondo processo di idealizzazione è strettamente legato alla scoperta delle grandezze incommensurabili. Come abbiamo visto, ci si era accorti che un segmento non può essere costituito da una sequenza finita di punti materiali come pretendevano i pitagorici. D'altra parte se un segmento contiene quanti punti si vuole, allora tali punti devono necessariamente avere lunghezza nulla. Infatti, se per assurdo tutti i punti avessero grandezza l allora il segmento in questione dovrebbe avere lunghezza pari alla somma di infinite volte l e cioè dovrebbe avere lunghezza infinita. In conclusione:

> i punti devono essere enti senza grandezza,[10]

e, per analoghe considerazioni,

> le linee devono essere enti senza larghezza
>
> le superfici enti senza spessore.

La idealizzazione degli enti geometrici assume allora un aspetto radicale, in quanto nel mondo reale ogni cosa ha lunghezza, larghezza e spessore. Naturalmente anche tutti gli altri enti geometrici come la sfera, il cubo, il cilindro sono astratti: non sarà mai possibile trovare in natura o costruire un corpo perfettamente sferico. Ma mentre il supporre l'esistenza di un corpo perfettamente sferico non sembra crearci grandi problemi, il supporre l'esi-

---

[10] Non è detto che i punti debbano essere assunti come concetti primitivi. In proposito si veda il mio articolo alla fine del capitolo.

stenza di qualcosa, il punto, che sia senza dimensioni è in completo contrasto con la concezione che abbiamo della materia. Siamo in presenza di un più alto livello di astrazione. In definitiva - la concezione del punto come ente senza dimensioni appariva come logica conseguenza della scoperta degli incommensurabili - i punti (più in generale, le linee, le superfici) non appartengono al mondo reale.

La conclusione a cui si doveva allora necessariamente pervenire era che:

si può avere conoscenza solo del mondo delle idee.

Il mondo percepito attraverso i sensi è qualcosa di illusorio al quale spetta solo il compito di "assomigliare" al mondo delle idee, così come un granello di sabbia può solo assomigliare ad un punto, non essere un punto. Un tale punto di vista assumerà la sua forma più compiuta nelle teorie filosofiche di Platone e sarà una delle cause dello scarsissimo sviluppo della fisica presso i Greci. Vediamo cosa dice Platone:

> *"I geometri si servono di figure visibili e ragionano su di esse, ma non ad esse pensando, bensì a ciò di cui quelle sono immagini, ragionando sul quadrato in sé e sulla diagonale in sé, e non su quella che disegnano. Lo stesso si dica per tutte le figure, che essi modellano e disegnano, di cui si servono come immagini (a guisa di ombre e di immagini riflesse nelle acque) cercando di vedere i veri enti che non si possono vedere se non col pensiero"* (Platone).

Quindi è il mondo reale ad essere una immagine (sbiadita) del mondo delle idee e non il contrario, come saremmo portati a pensare ora. In altre parole, attualmente se una teoria non descrive bene la realtà, allora è la teoria che viene considerata non adeguata e non certo la realtà.

Ma come essere sicuri della validità di una asserzione geometrica dal momento che, non riferendosi al mondo reale, la sua verifica non può essere sperimentale? I greci osservarono che di alcune asserzioni si ha una intuizione talmente immediata che tutti sono d'accordo sulla loro validità. Un esempio di tale tipo è *"per due punti distinti passa una sola retta"*. Vi sono però asserzioni, come il teorema di Pitagora, la cui validità non sembra essere altrettanto immediata. L'unica possibilità appariva allora quella di ricavare, tramite opportuni ragionamenti, le asserzioni più complesse da un gruppo prefissato di asserzioni sulla cui validità ci fosse un accordo generale: tali asserzioni venivano chiamate in alcuni casi postulati, in altri assiomi. Nasce in tale modo il metodo assiomatico che troverà negli elementi di Euclide la sua applicazione più bella e completa.

8. Gli elementi di Euclide.

Gli "Elementi" di Euclide rappresentano forse la tappa più importante dello sviluppo del pensiero scientifico moderno. L'opera, che consiste in 13 libri, fu composta da Euclide verso il 300 a.C. e rappresenta una organica esposizione di buona parte della matematica preesistente. Di questi libri i primi sei sono dedicati alla geometria piana, il settimo, l'ottavo, il nono ed il decimo sono di natura aritmetica, i rimanenti trattano di geometria solida.

Si inizia poi con le seguenti quattro definizioni.

Definizioni

1. *Il punto è ciò che non ha parti*
2. *Una linea è una lunghezza senza larghezza*
3. *Estremi di una linea sono punti*
4. *Linea retta è quella che giace ugualmente rispetto ai suoi punti.*

Una critica che a volte è stata fatta a tali definizioni è che rimandano il concetto da definire ad altri concetti che non si sono definiti. Infatti per definire il punto si ricorre alla nozione di "parte", per definire le linee si utilizzano le nozioni di "lunghezza" e "larghezza", in 3 si usa la parola "estremi". Infine è alquanto oscuro che cosa significhi "giacere ugualmente". Invece attualmente una definizione viene fatta solo in funzione di nozioni già definite. Ad esempio data una struttura algebrica, un suo elemento e viene chiamato "elemento neutro" se risulta $x \cdot e = x$ e $e \cdot x = x$ per ogni elemento x. Tale definizione è ragionevole poiché utilizza solo la nozione di prodotto il quale, per il fatto che partiamo da una struttura algebrica data, risulta essere stata già definita.

Tali critiche sono però ingiustificate perché per i greci le definizioni avevano un significato completamente diverso da quello attuale. Il loro ruolo (in un certo senso precedente ed esterno alla elaborazione scientifica) era quello di indicare, in qualche modo, enti che si riteneva avessero una esistenza propria e di cui ogni uomo ha una idea chiara. Prima di cominciare una trattazione scientifica di tali oggetti era infatti necessario, per potersi capire, essere sicuri che si stava parlando delle stesse cose. Ad esempio la Definizione 3 ci serve per capire che il termine "linea retta" significava quello che ora chiamiamo "segmento" e non ciò che oggi si intende per retta illimitata.[11]

In proposito si parla anche di *"definizioni reali"* in quanto sono simili alle descrizioni-indicazioni che facciamo a volte di un oggetto esistente e conosciuto sia da noi che dalla persona con cui parliamo. Così "il punto è ciò che non ha parti" è una definizione allo stesso modo per cui lo è, ad esempio, "la penna di cui intendo parlare è quella posata sul tavolo". A volte si parla anche di "definizioni ostensive" per esprimere il fatto che esse servono solo ad indicare, a mostrare l'oggetto di cui si parla.

Il punto di vista attuale, come vedremo quando esamineremo il metodo assiomatico, è totalmente diverso. Non si definisce ciascun ente isolatamente ma una intera classe di strutture ciascuna costituita da elementi matematici. Ad esempio non si propone una definizione di punto e di retta, piuttosto si definiscono delle strutture chiamate "spazi geometrici" i cui elementi base sono punti e rette. Oppure, le definizioni hanno il ruolo di indicare particolari elementi di una struttura.[12]

Negli Elementi di Euclide abbiamo poi una serie di *nozioni comuni* che riguardano l'uguaglianza o l'ordinamento.

Nozioni comuni

1. *cose che sono uguali alla stessa cosa sono uguali tra loro*
2. *cose che coincidono tra loro sono uguali*
3. *se cose uguali sono addizionate a cose uguali, le totalità sono uguali.*
4. *se cose uguali sono addizionate a cose diseguali le totalità sono disuguali*
5. *il tutto è maggiore della parte.*

[11] D'altra parte, il linguaggio usato ormai universalmente nei corsi di matematica è quello della teoria degli insiemi. Tuttavia gli insiemi non vengono mai definiti. Chi insegna cerca di dire che cosa è un insieme usando parole come "collezione", "gruppo", "aggregato" oppure facendo esempi. In questo si comporta in modo analogo a quello di Euclide cercando solo di mostrare quello di cui si parla.

[12] Ad esempio, in algebra, abbiamo sia la definizione di che cosa si debba intendere, ad esempio, per monoide, sia, dato un monoide, la definizione di che cosa sia un elemento neutro.

La proprietà simmetrica per cui se $A = B$ allora $B = A$ viene data per scontata e non esplicitata. Il primo assioma dice che la uguaglianza è una relazione transitiva. Infatti afferma che da $A = C$ e $B = C$ (equivalentemente $C = B$) segue $A = B$. Il secondo esprime la proprietà riflessiva. Pertanto i primi due assiomi, insieme alla proprietà simmetrica, ci dicono che = è una relazione di equivalenza. Il terzo assioma dice che l'eguaglianza è quella che attualmente viene chiamata una congruenza, cioè una equivalenza compatibile con la struttura matematica che si considera (in questo caso la struttura additiva).[13] L'ultimo assioma esprime la compatibilità dell'operazione + con la relazione d'ordine.

Infine seguono i cinque famosi postulati. Il quinto è noto come "assioma delle parallele" e, come vedremo, giocherà un ruolo importante nello sviluppo del pensiero matematico.

Postulati

I *Risulti postulato che si possa condurre una linea retta da una qualsiasi punto ad ogni altro punto.*

II *che una retta finita si possa prolungare continuamente in linea retta*

III *che si possa descrivere un cerchio con qualsiasi centro ed ogni distanza*

IV *che tutti gli angoli retti siano uguali tra loro*

V *che, se una retta venendo a cadere su due rette forma due angoli interni e dalla stessa parte minori di due retti, le due rette prolungate illimitatamente verranno ad incontrarsi da quella parte in cui sono gli angoli minori di due retti.*

Le nozioni comuni si differenziano dai postulati per il fatto di non appartenere esclusivamente alla geometria ma a tutte le scienze. Inoltre, in un certo senso, esse hanno un grado di certezza maggiore dei postulati. Da notare che sia le nozioni comuni che i postulati saranno chiamati dai matematici posteriori con il nome di *assiomi*. I matematici greci erano molto più prudenti poiché né il termine "nozione comune" né quello di "postulato"

[13] Da notare che l'esigenza di affermare esplicitamente che cose che coincidono sono uguali mostra che l'uguaglianza non coincide in generale con l'identità, cioè che possono esistere cose uguali ma non identiche. D'altra parte due triangoli vengono definiti uguali se hanno lati uguali. Quindi due triangoli possono essere uguali senza coincidere (in quanto situati in parti diverse del piano).

hanno pretese universali. Con il primo termine si indicava qualche cosa accettata da tutti i membri di una comunità. Con il secondo termine solo qualche cosa di cui si "chiede" (e da ciò deriva il termine "postulato") una accettazione al proprio interlocutore per poter rendere possibile la successiva trattazione.

Evitare infinito ed illimitato. Come abbiamo già osservato, il rifiuto dell'infinito attuale porta i matematici greci ad evitare accuratamente il ricorso ad enti infiniti o illimitati, cioè all'infinito attuale. Così al posto di quelle che per noi sono le rette illimitate i greci si riferiscono costantemente ai segmenti. La illimitatezza (attuale) della retta si traduce, nel secondo postulato, nella indefinita prolungabilità dei segmenti, e lo stesso quinto postulato si riferisce a prolungamenti di segmenti. Non si deve poi pensare che i segmenti venissero considerati come oggetti infiniti (insiemi infiniti di punti). Infatti se è vero che dagli assiomi di Euclide è possibile dedurre che in ogni segmento giacciono "quanti punti si vuole", per i matematici greci questo non significava che un segmento è un insieme infinito di punti. Piuttosto segmenti e punti erano enti completamente indipendenti tra i quali sussisteva o meno una relazione di "giacenza" da non confondersi in nessun modo con l'attuale relazione di appartenenza della teoria degli insiemi. Il fatto che, dato un segmento, se si trovano n punti giacenti in esso sia possibile trovarne anche $n+1$ non comporta l'accettazione dell'infinito attuale più di quanto lo comporti il fatto che, dato il numero intero n esista anche il numero $n+1$.

Carattere costruttivo dei postulati. Una ultima osservazione circa i postulati riguarda il loro carattere costruttivo. Si vede infatti chiaramente come essi siano corrispondenti ad operazioni che un disegnatore può eseguire, come il tracciare rette e cerchi. Fa eccezione forse il quinto postulato perché se la somma dei due angoli interni è minore di due retti solo per una quantità piccolissima, il verificare che le due rette in questione si incontrano comporta la necessità di tracciare segmenti più lunghi di quanto l'uomo è in grado di fare.

In definitiva potremmo anche dire che i postulati consentono di costruire, a partire da enti "a portata di mano" ancora enti "a portata di mano". Invece il quinto assioma sembra comportarsi in modo diverso in quanto si riferisce al fatto che un punto di incontro tra le due rette prima o poi verrà trovato anche se (per dirla in modo un po' rozzo) questo punto di incontro fosse tanto lontano da non potere effettivamente essere costruito. Per questo motivo, o forse anche per altri, il quinto postulato venne accettato malvolentieri ed i matematici, da Euclide fino alla prima metà dell'ottocento, cercarono costantemente di eliminarlo dimostrandolo a partire dagli altri assiomi. Si deve comunque sottolineare che non veniva messo in discussione il fatto che tale postulato fosse vero. Solo che, per questioni di correttezza scientifica e di eleganza sarebbe stato opportuno ricorrere solo a postulati totalmente evidenti.

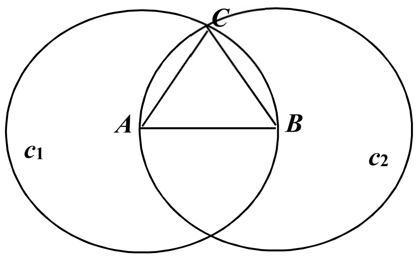
Il quinto postulato attualmente viene sostituito dal seguente:

> ***Postulato delle parallele:*** *data una retta r ed un punto P fuori da essa esiste al più una retta per P parallela ad r*

che si prova essere equivalente. Si tenga presente che l'esistenza della retta parallela può essere dimostrata a partire dai rimanenti postulati.

Concludiamo questo paragrafo con la prima delle proposizioni dimostrate da Euclide.

Proposizione 8.1. E' possibile, dato un segmento AB, costruire un triangolo equilatero di lato uguale ad AB.



Dim. Si tracci la circonferenza c_1 di centro A ed apertura AB e la circonferenza c_2 di centro B ed apertura BA e sia C il punto di incontro di tali circonferenze. Allora il triangolo ABC è equilatero. Infatti AC è uguale ad AB in quanto entrambi raggi di c_1 e CB è uguale ad AB in quanto raggi di c_2. $\square$

Come si vede si tratta di una dimostrazione semplice ed elegante …, purtroppo proprio in questa che è la prima delle dimostrazioni presente negli Elementi di Euclide si presenta un errore !

Proposizione 8.2. La dimostrazione della Proposizione 8.1 è sbagliata. Inoltre nel sistema di assiomi di Euclide non possono esistere dimostrazioni corrette di tale proposizione.

Dim. L'errore consiste nell'asserire, senza che sia stato dimostrato, che le due circonferenze si incontrano in un punto C. Per potere dimostrare questo fatto è necessario un postulato ulteriore, detto *postulato di continuità* che analizzeremo nel seguito. Nei suoi *Elementi* Euclide non pone in modo esplicito tale postulato anche se poi di fatto lo utilizza.

Naturalmente si potrebbe anche pensare di trovare una dimostrazione diversa che non usi la proprietà di continuità. Purtroppo una tale dimostrazione non può esistere. Infatti consideriamo un modello di geometria in cui i punti sono gli elementi del prodotto cartesiano $Q\times Q$, cioè le coppie di numeri razionali ed in cui una retta è l'insieme dei punti che soddisfano una data equazione di primo grado a coefficienti razionali. Chiamiamo *piano razionale* tale modello. Non è difficile controllare che nel piano razionale tutti i postulati di Euclide sono verificati e quindi lo sono anche tutti teoremi che seguono da tali postulati. Se allora si potesse dimostrare la Proposizione 8.1, tale proposizione dovrebbe essere verificata dal piano razionale. Ciò non è vero. Infatti consideriamo i punti $A = (-1,0)$ e $B = (1,0)$, e supponiamo che esista C tale che ABC sia equilatero. Allora tale punto dovrebbe essere del tipo $(0,q)$ e, per il teorema di Pitagora, q^2 dovrebbe essere uguale a 3. Non esistendo un razionale il cui quadrato sia 3, ciò è assurdo. □

9. Teoria delle grandezze omogenee (invece dei numeri reali)

Le nozioni di classe di grandezze omogenee e di proporzione tra grandezze omogenee si trovano nel libro V degli *Elementi* di Euclide e sembrano risalire a Eudosso di Cnido, vissuto pochi decenni prima di Euclide. Tali nozioni permettono ai greci di fare molte delle cose che attualmente vengono fatte ricorrendo ai numeri reali positivi. Negli *Elementi* non viene proposto un sistema di assiomi completo per la nozione di classe di grandezze omogenee e spesso Euclide utilizza proprietà dettate dall'intuizione. Invece noi proviamo a proporre il seguente sistema di assiomi.

Definizione 9.1. Una *classe di grandezze omogenee* è una struttura $(G, =, <, +)$, tale che

$A1$ + è una operazione commutativa ed associativa.

$A2$ per ogni $a\in G$ e per ogni $n\in N$ esiste $u\in G$ tale che $n\cdot u = a$

(assioma della divisibilità).
$A3$ $<$ è una relazione d'ordine stretto totale compatibile con +.

Naturalmente si suppone che l'eguaglianza = verifichi le proprietà elencate nelle nozioni comuni e che quindi sia una equivalenza compatibile con + e <. Due grandezze che appartengono alla stessa classe di grandezze omogenee si dicono *omogenee* tra loro.

La classe dei numeri naturali soddisfa $A1$ e $A3$ ma non $A2$. Esempi di classe di grandezze omogenee sono i seguenti:
- la classe dei razionali positivi
- la classe dei reali positivi

Un esempio che più esprime l'idea dei greci di classe di grandezze omogenee è dato dall'insieme dei pesi di una bilancia. Chiamiamo uguali due pesi A e B se posti sui due piatti della bilancia rimangono in equilibrio. Diciamo invece che A è maggiore di B se il piatto su cui si poggia A si abbassa. La definizione di somma di due pesi è ovvia. Se vogliamo che l'assioma della divisibilità sia verificato dobbiamo accettare che dato un peso abbiamo a disposizione anche tutti i suoi sottomultipli.

Oltre agli assiomi elencati sono importanti anche il *Postulato di Archimede* e l'*Assioma della continuità* che presentano un interesse notevole anche nella matematica moderna ed in particolare nella teoria dei campi ordinati.

Postulato di Archimede. Siano u e b due grandezze omogenee con $u < b$, allora esiste un intero m tale che $m \cdot u > b$.

Tutti gli esempi numerici precedenti verificano l'assioma di Archimede. Per dare una idea di un comportamento non archimedeo, sia R l'insieme dei numeri reali e consideriamo la classe R^R delle funzioni di R in R. Inoltre definiamo la somma di due funzioni e l'ordinamento $\leq$ tra funzioni in modo usuale. In tale modello è subito visto che l'assioma di Archimede non è verificato. Infatti sia f una funzione limitata e g una funzione non limitata. Allora tutti i multipli di f sono ancora limitati e quindi non può esistere un multiplo di f maggiore di g. Il fatto che tale tipo di struttura non è archimedea comporta che non risulta possibile misurare le funzioni utilizzando una funzione come unità di misura come invece è possibile fare per i segmenti.

Da notare che in questo esempio sono verificati tutti gli assiomi per le classi di grandezze omogenee tranne il fatto che l'ordinamento non è totale. Un esempio migliore di struttura non archimedea è quello dei razionali non standard che considereremo nel seguito.

L'assioma della continuità afferma, in un certo senso, che non esistono "buchi" in una classe di grandezze omogenee.[14] Esso comunque, pur essendo spesso utilizzato, non fu mai enunciato esplicitamente dai matematici greci. Bisogna aspettare Dedekind nel 1872, per avere una sua esplicita formulazione.

Definizione 9.2. Chiameremo *separati* due sottoinsiemi A e B di una classe di grandezze omogenee se ogni elemento di A è minore di ogni elemento di B. Un *elemento di separazione* è un elemento che è maggiorante di A e minorante di B.

Assioma di continuità (o di completezza). Ogni coppia A e B di insiemi separati ammette un elemento di separazione.

Si osservi che tale assioma non è verificato dalla classe $Q^+ = \{q \in Q : q > 0\}$ dei numeri razionali positivi. Infatti è possibile provare che le due classi $A = \{r \in Q^+ \mid r^2 < 2\}$ e $B = \{r \in Q^+ \mid r^2 > 2\}$ sono separate ma che non esiste nessun razionale q che sia elemento di separazione.

Classe di grandezze omogenee e processo di misurazione.
Il sistema di assiomi che abbiamo elencato ha lo scopo di rendere possibile una misurazione di una grandezza b rispetto ad una unità di misura u qualunque. Ad esempio assumiamo che u sia un regolo lungo un decimetro e che vogliamo misurare un segmento b. Allora la presenza di una operazione di somma consente di effettuare multipli $1 \cdot u$, $2 \cdot u$, ..., $m \cdot u$ di u cioè di riportare più volte consecutivamente u. Allo stesso momento, utilizzando il fatto che in una classe di grandezze omogenee è definita una uguaglianza ed un ordine stretto, possiamo man mano verificare se il multiplo $m \cdot u$ sia uguale o minore a b. Ad un certo momento è possibile che si ottenga un multiplo $m \cdot u$ di u coincidente proprio

[14] Abbiamo già accennato all'esigenza di un postulato di continuità nell'esaminare la dimostrazione del primo dei teoremi dimostrati negli *Elementi*.

con b. In questo caso "fortunato" diremo che *m è la misura di b rispetto all'unità di misura u*. Se ciò non accade, allora per il Postulato di Archimede ad un certo punto si finisce comunque col *superare b*. Esiste cioè un numero m tale che $(m\text{-}1)\cdot u<b<m\cdot u$. Allora possiamo assumere i numeri m-1 ed m rispettivamente come *misura per difetto* e *misura per eccesso* di b rispetto a u. Se poi vogliamo ottenere una misura migliore, possiamo sostituire u con un suo sottomultiplo ad esempio porre $u' = u/10$ uguale al centimetro, cosa questa che l'assioma di divisibilità consente di fare. In questo caso, se esiste un intero m tale che $m\cdot u' = b$, cioè tale che $(m/10)\cdot u = b$ diciamo che *la misura di b rispetto a u è il numero razionale* $m/10$. Altrimenti, detto m un intero tale che $(m\text{-}1)\cdot u'<b<m\cdot u'$ cioè tale che $((m\text{-}1)/10)\cdot u<b<(m/10)\cdot u$ diremo che b ha *come misura per difetto* il razionale $(m\text{-}1)/10$ e *misura per eccesso* il razionale $m/10$.[15] Se si vogliono misure più precise si procede poi dividendo il segmento u' in ulteriori 10 parti ottenendo un regolo di un millimetro e così via.

Non cambia il procedimento se invece che misure di segmenti si tratta di pesare oggetti. Se b è l'oggetto da pesare ed u (*il grammo*) un oggetto che funge da una unità di misura, allora si procede in questo modo. Si mette b in un piatto B ed ovviamente B si abbassa. Nel frattempo sull'altro piatto A si mettono man mano dei pesi uguali ad u. Se dopo avere messo m pesi uguali ad u la bilancia si equilibra, allora possiamo concludere che la misura di b rispetto ad u è esattamente m. In caso contrario ad un certo momento (assioma di Archimede) il piatto B si alza e possiamo concludere che il numero m dei pesi in A è una misura per eccesso di b, mentre m-1 ne è una misura per difetto. Utilizzando sottomultipli di u si possono ottenere delle pesate più precise.

Per fare invece un esempio del ruolo giocato dall'assioma della continuità, consideriamo il seguente problema:

data una circonferenza C trovare un segmento la cui lunghezza sia uguale a quella di C.

Allora posso considerare l'insieme A dei segmenti che si ottengono inscrivendo una poligonale in C e poi "raddrizzandola" in un segmento. In altre parole A è l'insieme dei segmenti la cui lunghezza è uguale alla lunghezza di una poligonale inscritta. Defi-

[15] Quest'ultimo passaggio naturalmente non poteva essere esplicitato dai greci per il fatto che essi non accettavano i razionali.

niamo inoltre l'insieme B come l'insieme dei segmenti la cui lunghezza si ottiene "raddrizzando" una poligonale circoscritta. Precisamente B è l'insieme dei segmenti la cui lunghezza è uguale alla lunghezza di una poligonale circoscritta. Allora A e B sono due classi separate e pertanto, per l'assioma di continuità, esiste un segmento che separa tali classi. Tale segmento rappresenta il segmento la cui la lunghezza è uguale a quella della circonferenza.

10. La teoria delle proporzioni (invece delle operazioni)

Abbiamo già parlato di incommensurabilità quando abbiamo provato che il lato e la diagonale del quadrato non sono commensurabili. Ci siamo riferiti ai segmenti ed abbiamo utilizzato l'idea intuitiva di somma di due segmenti intesa come mettere uno dopo l'altro i segmenti disponendoli sulla stessa linea ed anche l'idea che, n segmenti $a_1,\ldots a_n$ sono commensurabili se esiste un segmento unitario u tale che tali segmenti siano tutti multipli di u. In questo paragrafo riprendiamo tale nozione estendendola alle classi di grandezze omogenee.

Definizione 10.1. Due grandezze a e b sono *commensurabili* se esiste u tale $n{\cdot}u = a$ e $m{\cdot}u = b$ dove n ed m sono opportuni naturali.

In termini moderni noi diremmo che a e b sono commensurabili se $a = (m/n)b$, ma i Greci come abbiamo già osservato, non potevano fare riferimento ai razionali. Quello che è per noi il numero razionale m/n rappresentava per loro una relazione che doveva essere espressa, se si voleva mantenere il dovuto rigore, solo in termini di interi tramite l'uguaglianza $n \cdot a = m{\cdot}b$. Ad esempio Euclide nei suoi Elementi afferma che

> *"un rapporto è una sorta di relazione tra dimensioni di due grandezze della stessa specie."*

Proposizione 10.2. Due grandezze omogenee a e b sono *commensurabili* se e solo se esistono due interi n ed m tali che $n{\cdot}a = m{\cdot}b$.

Dim. Supponiamo che esistano due naturali n ed m tali che $n{\cdot}a = m{\cdot}b$. Sia u tale che $m{\cdot}u = a$ allora risulta che $mnu = nmu = mb$ e quindi $b = nu$. Ciò significa che sia a che b sono multipli di u.

Viceversa sia u tale che, per opportuni n ed m, $mu = a$ e $nu = b$, allora $nmu = na$ e $mnu = mb$ e quindi $na = mb$. □

Si passa poi alla definizione della nozione di proporzionalità tra quattro grandezze. Anche in questo caso si utilizza solo la nozione di multiplo e non quella di divisione. Ora se attualmente volessimo esprimere il fatto che quattro grandezze a, b, c, d (che in generale possono essere non razionali) sono in proporzione, cioè che il numero reale $a : b$ è uguale al numero reale $c{:}d$, potremmo farlo ricorrendo solo ai razionali dicendo che
- un razionale n/m è minore di $a{:}b$ se e solo se è minore di $c{:}d$
- un razionale n/m è maggiore di $a{:}b$ se e solo se è maggiore di $c{:}d$.

D'altra parte possiamo riscrivere tali equivalenze facendo uso solo dei numeri interi dicendo che:
- $nb \leq ma$ se e solo se $nd \leq mc$
- $nb \geq ma$ se e solo se $nd \geq mc$.

Ciò suggerisce la seguente definizione dove diciamo che due o più grandezze sono *omogenee* tra loro se appartengono alla stessa classe di grandezze omogenee.

Definizione 10.3. Siano a, b, c, d quattro grandezze con a omogeneo a b e c omogeneo a d. Diremo che tali grandezze *sono in proporzione* e scriveremo $a{:}b = c{:}d$ se per ogni coppia di interi n, m risulta che

$$nb \leq ma \Leftrightarrow nd \leq mc \quad \text{e} \quad nb \geq ma \Leftrightarrow nd \geq mc.$$

Si noti che non è necessario che le quattro grandezze siano omogenee tra loro, è sufficiente assumere che a sia omogenea a b e c a d; cioè che a e b appartengano ad una stessa classe di grandezze omogenee e c e d ad un'altra classe di grandezze omogenee. Ad esempio è possibile parlare di proporzione anche nel caso in cui a e b sono lunghezze e c e d aree.

Nota sul simbolo di uguaglianza. Da osservare che l'uso dell'uguaglianza nell'espressione $a{:}b = c{:}d$ non deve far ritenere che i Greci ritenessero che $a{:}b$ e e $c{:}d$ fossero "oggetti" uguali. Infatti a e b non sono da considerare numeri reali di cui si effettua la divisione e quindi $a{:}b$ da solo non denota niente. La equazione $a{:}b = c{:}d$ era considerata come un modo per indicare una relazione tra quattro grandezze e non una uguaglianza tra due. Il segno di uguaglianza può verificarsi soltanto nel caso di coppie di grandezze commensurabili che sono le uniche che ammettono multipli comuni. Tuttavia sembra che Euclide consideri una scrittura del genere anche come una sorta di affermazione per cui la coppia (a,b) ha qualcosa in comune con la coppia (c,d) (si veda la nota successiva).

Proposizione 10.4. Per l'uguaglianza dei rapporti valgono le proprietà riflessiva, simmetrica e transitiva (cioè è una relazione di equivalenza tra coppie).[16]

Dim. La proprietà riflessiva e quella simmetrica sono evidenti. Per provare la proprietà transitiva siano a, b, c, d, e, f sei grandezze diverse tra loro, vogliamo provare che

$$a : b = c : d \quad \text{e} \quad c : d = e : f \;\Rightarrow\; a : b = e : f.$$

A tale scopo supponiamo che $ma \leq nb$, allora per la prima proporzione $mc \leq nd$. D'altra parte per la seconda proporzione si ha che da $mc \leq nd$ segue che $me \leq nf$. Quindi per ogni m, n in N abbiamo che $ma \leq nb \Rightarrow me \leq nf$. Similmente si prova che $ma \geq nb \Rightarrow m \cdot e \geq n \cdot f$ e ciò prova che $a{:}b = e{:}f$. □

Da notare che l'insieme dei numeri naturali costituisce una classe di grandezze omogenee, quindi possiamo scrivere una proporzione del tipo $a : b = p : q$ con p e q numeri naturali.
Concludiamo enunciando il teorema di esistenza del quarto proporzionale.

[16] Se si assume il punto di vista della matematica moderna, il fatto che valgano le proprietà riflessiva, simmetrica e transitiva induce a considerare una nuova classe di oggetti che si possono costruire al modo seguente. Data una classe G di grandezze omogenee definiamo in $G \times G$ la relazione $\equiv$ ponendo $(a,b) \equiv (c,d)$ se a, b, c, d sono in proporzione. Poiché una tale relazione è di equivalenza, ripartisce $G \times G$ in classi di equivalenza. Chiamiamo *"rapporto"* una classe completa di equivalenza. In questo modo è lecito dire che le due classi a e b hanno lo stesso rapporto per dire che appartengono alla stessa classe di equivalenza. Da questo punto di vista il simbolo di uguaglianza nell'espressione $a{:}b = c{:}d$ riacquista il suo significato usuale di identità.

Teorema 10.5. (*Esistenza del quarto proporzionale*). Siano G e G' classi di grandezze omogenee, a, $b \in G$ e $c \in G'$, allora esiste $x \in G'$ tale che

$$a : b = c : x. \tag{10.1}$$

Tale teorema equivale all'affermazione che in una classe di grandezze omogenee la moltiplicazione e la divisione sono operazioni ovunque definite. Infatti, supponendo che $G = G'$, date b e c ponendo $a = 1$ in (10.1), otteniamo che $1:b = c: x$ cioè (con il linguaggio attuale della teoria dei numeri reali) che $x = b \cdot c$. D'altra parte, dati b e c, se si pone $b = 1$ allora (10.1) diviene $a :1 = c:x$ che equivale a dire che $c = ax$.

11. Misura, equiscomponibilità, equicompletabilità

La teoria delle proporzioni costituiva uno strumento in qualche modo sostitutivo della teoria dei numeri reali. Tuttavia la "misurazione" di una figura geometrica (calcolo dell'area, volume, perimetro) non era intesa, come oggi, come assegnazione di un numero reale a tale figura. Per gli antichi greci essa consisteva nel confronto tra le grandezze di figure geometriche. Questo confronto consisteva nel constatare una uguaglianza o una proporzione tra le estensioni di due figure. Ad esempio, invece di dire che l'area del triangolo è uguale alla misura della base per l'altezza diviso due, si diceva che un triangolo ha la stessa estensione di un rettangolo con la stessa base e metà dell'altezza. Equivalentemente, si diceva che la sua estensione sta a quella di un rettangolo con stessa base e stessa altezza come 1 sta a 2.

Per quello che riguarda l'uguaglianza delle estensioni, una nozione che spesso veniva utilizzata era quella di "equiscomponibilità". Detto in termini intuitivi, tale nozione può essere definita al modo seguente.

> *Due figure geometriche F e $\underline{F}$ si dicono equiscomponibili se è possibile*
> *Tagliare* F *nelle figure* $X_1,\dots,X_n$,
>
> *Spostare* *tali figure in modo da ottenere le figure* $\underline{X}_1,\dots,\underline{X}_n$
>
> *Ricomporre* $\underline{X}_1,\dots,\underline{X}_n$ *in modo da ottenere* $\underline{F}$.

In altre parole due figure sono equiscomponibili se sono la "somma" di figure uguali. Naturalmente è necessario precisare che cosa si intenda per "figura geometrica", "tagliare", "spostare", "ricomporre". per gli antichi greci tali nozioni erano considerate primitive e quindi non definibili. In questo libro ci riferiamo invece una interpretazione in termini matematici attuali. Ad esempio accettiamo che per piano euclideo si intenda R^2 e che per figura geometrica si intenda un qualunque insieme di punti (cosa alquanto discutibile come vedremo nel seguito). Allora la nozione di "tagliare" può essere rappresentata dalla nozione insiemistica di partizione.

Definizione 11.1. Se X è un insieme chiamiamo *partizione finita* di X, una classe finita $X_1,\dots,X_n$ di insiemi non vuoti tali che

- $X = X_1 \cup \dots \cup X_n$
- per ogni i e j con $i \neq j$, $X_i \cap X_j = \emptyset$.

Per quanto riguarda la nozione "spostare" si può ricorrere alla nozione di isometria.

Definizione 11.2. Chiamiamo *isometria* del piano euclideo E ogni funzione $i : R^2 \to R^2$ che conserva le distanze, cioè tale che

$$d(x,y) = d(i(x),i(y)).$$

Diciamo che i *sposta* un insieme di punti A nell'insieme di punti B se $i(A) = B$. Due figure geometriche A e B si dicono *isometriche* o *uguali* se esiste una isometria che sposta A in B.

Esempi tipici di isometrie nel piano sono le traslazioni, le rotazioni, ed i ribaltamenti rispetto ad un asse. E' evidente che la relazione di isometria tra insiemi è una relazione di equivalenza.

Possiamo ora dare la nozione di equiscomponibilità in modo più preciso.

Definizione 11.3. Due figure geometriche F e $\underline{F}$ si dicono *equiscomponibili* se ammettono rispettivamente due partizioni $X_1,\dots,X_n$, e $\underline{X}_1,\dots,\underline{X}_n$, tali che, per ogni i, X_i è isometrico a $\underline{X}_i$.

Da notare che due figure isometriche risultano essere equiscomponibili. La seguente proposizione mostra perché l'equiscomponibilità è uno strumento per il confronto tra le estensioni delle figure. Da notare che ci riferiamo a "misure" intese come funzioni finitamente additive ed invarianti per isometrie.

Teorema 11.4. Se due figure F e $\underline{F}$ sono equiscomponibili allora hanno la stessa misura.[17]

Dim. Se μ è la funzione-misura allora, con riferimento alle notazioni della definizione 10.3,

$$\begin{aligned} \mu(F) &= \mu(X_1 \cup \ldots \cup X_n) = \mu(X_1) + \ldots + \mu(,X_n) \\ &= \mu(\underline{X}_1) + \ldots + \mu(,\underline{X}_n) = \mu(\underline{X}_1 \cup \ldots \cup \underline{X}_n) \\ &= \mu(\underline{F}). \end{aligned}$$
□

Abbiamo visto che due figure sono equiscomponibili se sono “somma” di pezzi uguali. Possiamo anche considerare la possibilità che due figure siano la “differenza” tra figure uguali.

Definizione 11.5. Due figure si dicono *equicompletabili* se aggiungendo ad esse parti a due a due uguali si ottengono figure uguali.

La dimostrazione della seguente proposizione è evidente.

Proposizione 11.6. Due figure equicompletabili hanno la stessa misura.

12. Esempi di calcolo delle aree

Due “popolari” applicazioni del teorema 11.4 permettono il calcolo dell’area di un parallelogramma e quella di un triangolo.

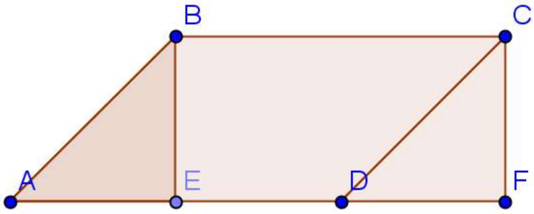
Teorema 12.1. Dato un parallelogramma fissiamo un lato come base, allora il parallelogramma è equiscomponibile ad un rettangolo con la stessa base e la stessa altezza.

Dim. Consideriamo un parallelogramma $ABCD$ e supponiamo che la base AD sia il lato maggiore. Riferendoci alla figura disegnata, proiettiamo i punti B e C sulla retta AD ottenendo i punti

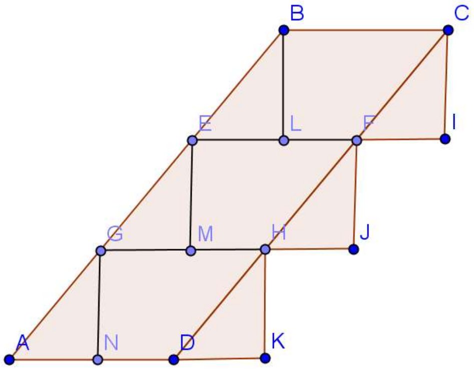
[17] Tale proposizione presuppone che la scomposizione delle due figure avvenga tramite pezzi che siano “misurabili”, cioè dotati di area. Senza tale ipotesi la proposizione non vale (si veda il paradosso di Vitali nel volume secondo). D’altra parte ai matematici greci appariva scontato che tutte le figure fossero misurabili, sia per la tipologia molto limitata delle figure considerate sia perché solo una impostazione insiemistica della geometria permette di trovare esempi di figure non misurabili.

E ed *F*, rispettivamente. Essendo *AE* minore di *AB* sarà anche minore di *AD* e quindi *AE* è interno al parallelogramma. I due triangoli *ABE* e *DCF* sono uguali. Per ottenere la decomposizione cercata basta allora tagliare il triangolo *ABE* e traslarlo nella posizione *DCF*. Si ottiene in tale modo il rettangolo *EBCF* avente la stessa base e la stessa altezza del parallelogramma.[18]

Nel caso in cui *AD* sia il lato minore, basta tagliare il parallelogramma in *n* strisce parallele ad *AD* con un lato uguale ad *AD* e l'altro minore di *AD*. Ciascuno di tali parallelogrammi rientra



nel caso che abbiamo già affrontato e quindi può essere ridotto ad un rettangolo di base *AD*.[19] Sovrapponendo i rettangoli così ottenuti otteniamo il rettangolo cercato.



□

Questo teorema corrisponde alla ben nota regola:

---

[18] In altre parole si taglia il triangolo $A_1A_2A_3$ che è a sinistra e lo si sposta a destra in modo da costruire il rettangolo.

[19] Per ridurre il numero di tagli conviene prendere il primo rettangolo, ad esempio quello di base *BC*, con altezza uguale ad *AD* e poi ripetere alla parte rimanente del parallelogramma la stessa operazione fino a quando la parte rimanente non diviene di altezza minore di *AD*.

L'area del parallelogramma è uguale alla base per l'altezza.

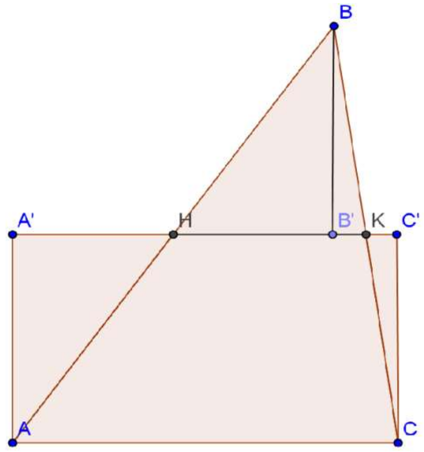
Ecco un'altra famosa utilizzazione della equiscomponibilità.

Teorema 12.2. Ogni triangolo è equiscomponibile ad un rettangolo di uguale base e metà altezza (qualunque sia il lato scelto come base).

H

Dim. Dato un triangolo di vertici *A*, *B*, *C* fissiamo come base il segmento *AC* e supponiamo che i due angoli in *A* e *C* siano acuti.. Procediamo poi al modo seguente:

1. costruiamo il punto di mezzo *H* del lato *AB*
2. costruiamo il punto di mezzo *K* del lato *BC*
3. tracciamo il rettangolo *A'C'AC*,
4. tracciamo l'altezza *BB'* del triangolo *HBK*.

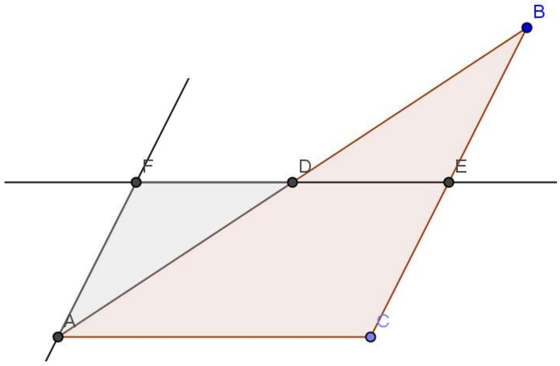


Stante l'ipotesi sugli angoli tale altezza è interna al triangolo. L'equiscomponibilità tra il triangolo *ABC* ed il rettangolo *A'C'AC* segue dal fatto che tali figure hanno il quadrilatero *AHKC* in comune, il triangolo *HBB'* è uguale a *AA'H* ed il triangolo *B'BK* è uguale a *KC'C*.[20] Tuttavia tale procedimento non può essere applicato nel caso in cui uno degli angoli in *A* ed in *C* sia ottuso.

[20] Possiamo visualizzare tale equiscomponibilità immaginando di tagliare *HBK* nei due triangolini e di ruotare tali triangolini intorno ai punti *H* e *K*.

Invece la seguente procedura funziona in tutti i casi.

1. Si costruisce il parallelogramma *AFEC* tracciando la retta per



i punti medi *D* ed *E* dei lati *AB* e *CB* e poi la parallela per *A* alla retta *CB*. Poiché il triangolo *ADF* risulta essere uguale al triangolo *DEB*, tale parallelogramma risulta essere equiscomponibile al triangolo *ABC.*
2. Si trasforma tale parallelogramma in un retangolo in accordo con la procedura del teorema precedente. □

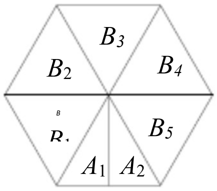
Il teorema ora dimostrato corrisponde alla ben nota regola:

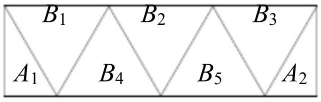
L'area del triangolo è uguale alla base per l'altezza diviso due.

Problema. Dimostrare che un trapezio è equiscomponibile ad un triangolo che ha come base la somma delle basi e la stessa altezza. Ricavarne la formula dell'area del trapezio.

Teorema 12.3. Un poligono regolare è equiscomponibile ad un rettangolo che ha come base la metà del perimetro e come altezza l'apotema.

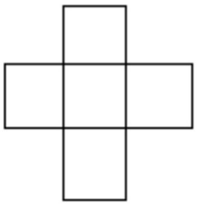
Dim. Riferendoci ad esempio ad un esagono, è sufficiente considerare la seguente figura:





in cui si mostra che l'esagono si può scomporre in 7 triangoli che ricomposti formano un rettangolo. □

Attualmente questo teorema si esprime dicendo che:
- *l'area di un poligono regolare è uguale al perimetro per l'apotema diviso due.*



Problema (difficile). Consideriamo una croce costituita da cinque quadrati. Provare che è equiscomponibile ad un quadrato Q (suggerimento: Se ogni quadratino ha lato 1 allora Q deve avere area 5. Quindi il suo lato elevato al quadrato deve dare 5. Per il teorema di Pitagora tale lato deve essere l'ipotenusa di un triangolo rettangolo di cateti …Pertanto il suo lato potrebbe essere il segmento …).

Un esempio di applicazione della nozione di equicompletabilità, è dato durante la dimostrazione del Teorema 3.1

13. L'equiscomponibilità è un metodo universale

Poniamoci ora il problema se il teorema 11.4 si possa invertire, cioè se tutte le volte che due figure hanno la stessa area allora sono equiscomponibili. La risposta è positiva, almeno per quanto riguarda le figure a contorno rettilineo, come è stato dimostrato da Bolyai nel 1832 e dal matematico dilettante Gerwin nel 1833. Questo comporta che il metodo dell'equiscomponibilità per il calcolo delle aree è universale. Per effettuare la dimostrazione per prima cosa dimostriamo che l'equiscomponibilità è una relazione di equivalenza. Tale proprietà segue sostanzialmente dal fatto che l'insieme delle isometrie costituisce un gruppo.

Proposizione 13.1. L'equiscomponibilità è una relazione di equivalenza, cioè è riflessiva, simmetrica e transitiva.

Dim. La proprietà riflessiva è ovvia. La proprietà simmetrica è una immediata conseguenza del fatto che ogni movimento ammette movimento inverso.

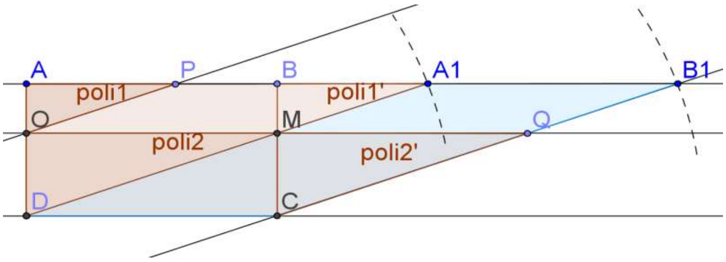
Per provare la transitività supponiamo che A sia equiscomponibile a B e che B sia equiscomponibile a C. Allora A può essere ripartito nelle figure $A_1,\ldots,A_n$ le quali opportunamente spostate tramite i movimenti $f_1,\ldots,f_n$ definiscono delle figure $\underline{A}_1 = f_1(A_1),\ldots, \underline{A}_n = f_n(A_n)$ tali che $B = \underline{A}_1\cup\ldots\cup\underline{A}_n$. Inoltre B può essere ripartito nelle figure $B_1,\ldots,B_m$ che opportunamente spostate tramite i movimenti $g_1,\ldots,g_m$ definiscono delle figure $\underline{B}_1 = g_1(B_1)$, ..., $\underline{B}_m = g_m(B_m)$ tali che $C = \underline{B}_1\cup\ldots\cup\underline{B}_m$. Allora:

- le figure $A_{ij} = f_i^{-1}(\underline{A}_i\cap B_j)$ ottenute al variare di i e j costituiscono una partizione di A,
- le figure $C_{ij} = g_j(f_i(A_{ij}))$ costituiscono una partizione di C.

Poiché il composto di due isometrie è una isometria, A_{ij} è uguale a C_{ij} e ciò prova che A e C sono equiscomponibili. □

Proposizione 13.2. Sia $ABCD$ un rettangolo, allora comunque si fissi un segmento EF esiste un rettangolo di base uguale ad EF equiscomponibile ad $ABCD$.

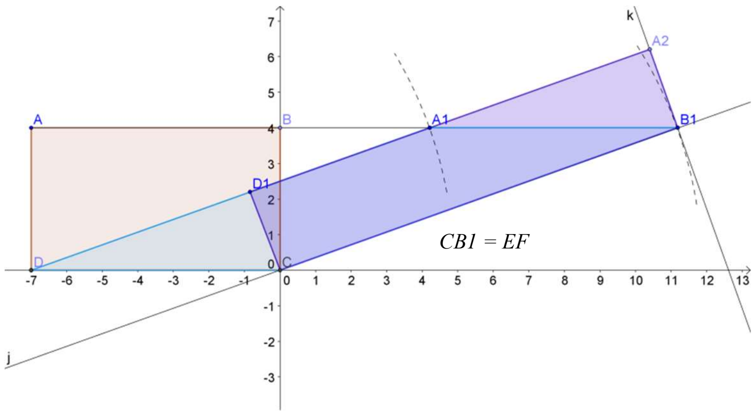
Dim. Supponiamo EF maggiore di CB, allora, tenendo ferma la base DC “deformiamo” il rettangolo $ABCD$ in un parallelo-



gramma DCA_1B_1 con A_1 e B_1 appartenenti alla retta AB ed in modo che il lato CB_1 sia uguale ad EF.[21] Successivamente “raddrizziamo” tale parallelogramma tenendo fisso il lato CB_1 e lasciando immutata l’altezza rispetto CB_1. Si ottiene un rettangolo

[21] Nella prima figura ciò è ottenuto tagliando il rettangolo tramite la retta per D parallela a CB_1 e spostando il triangolo poli2 in poli2’. Successivamente tagliando tramite la retta per O parallela a CB_1 e spostando il triangolo *poli*$_1$ in *poli*$_1$’. Infine il parallelogramma OPA_1M viene spostato nel parallelogramma MA_1B_1Q.

$CB_1D_1A_2$. Essendo $ABCD$ equiscomponibile al parallelogramma A_1B_1DC ed essendo tale parallelogramma equiscomponibile con il rettangolo $CB_1D_1A_2$, $ABCD$ è equiscomponibile con $CB_1D_1A_2$.



$ABCD \to A_1B_1DC \to CB_1D_1A_2$

Nel caso in cui EF sia minore di CB, possiamo dividere CB in n parti uguali in modo che CB/n sia minore di EF. Tagliamo poi $ABCD$ in n corrispondenti rettangoli di lato uguale a CB/n. Possiamo ora applicare a ciascuno di tali rettangoli la procedura sopra esposta ottenendo n rettangoli con un lato pari a EF. Sovrapponendo tali rettangoli otteniamo un rettangolo di lato uguale ad EF che risulta essere il rettangolo cercato. □

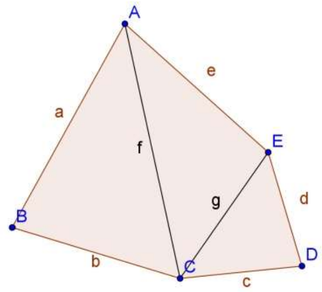
Corollario 13.3. Dati n rettangoli ed un segmento EF possiamo tagliare in pezzi tali rettangoli e ricomporli in un unico rettangolo di base EF.

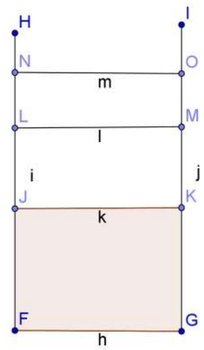
Dim. Abbiamo visto che ciascuno degli rettangoli è equiscomponibile ad un rettangolo di base EF. Sovrapponendo tali rettangoli otteniamo il rettangolo cercato. □

Teorema 13.4. Dato un segmento EF, ogni figura poligonale Q è equiscomponibile con un rettangolo di base uguale ad EF.

Dim. Tagliamo la figura poligonale Q in un numero finito di triangoli $T_1,\ldots,T_n$. Tali triangoli sono equiscomponibili a rettangoli $R_1,\ldots,R_n$ che a loro volta, come abbiamo ora visto, possono essere trasformati in un unico rettangolo di base uguale ad FG.

E’ sufficiente allora incollare tali rettangoli lungo le basi per ot-





tenere un rettangolo equiscomponibile a Q (nella figura Q è un pentagono che si divide in tre triangoli). □

Teorema 13.5. Due figure poligonali che sono di uguale misura sono equiscomponibili.[22]

Dim. Se due poligonali Q_1 e Q_2 hanno misura uguale allora, fissato un segmento AB esiste un rettangolo R_1 di base uguale ad AB equiscomponibile a Q_1 ed un rettangolo R_2 di base AB equiscomponibile a Q_2. Poiché Q_1 e Q_2 hanno per ipotesi la stessa area, R_1 ed R_2 hanno la stessa area e quindi sono uguali. In conclusione essendo Q_1 e Q_2 equiscomponibili a rettangoli uguali sono equiscomponibili tra loro. □

Tale teorema può essere completato al modo seguente.

Teorema 13.6. Siano Q_1 e Q_2 due figure poligonali, allora le seguenti asserzioni sono equivalenti.

i) Q_1 e Q_2 sono equicompletabili.
ii) Q_1 e Q_2 sono equiscomponibili.
iii) Q_1 e Q_2 hanno la stessa area.

[22] I matematici si sono posti il problema se un teorema simile non possa essere dimostrato anche nella geometria dello spazio. Tale problema fu incluso al terzo posto della famosa lista di problemi proposti da Hilbert nel 1900. La risposta è negativa: nel 1903 il matematico Dehn mostrò che esistono due tetraedri di uguale volume che non sono equiscomponibili.

Dim. Dobbiamo provare solo l'equivalenza tra *i*) e *ii*). Ora se Q_1 e Q_2 sono equicompletabili hanno la stessa area e quindi, per il teorema precedente, sono equiscomponibili. Se Q_1 e Q_2 sono equiscomponibili, detto Q un quadrato che le contiene entrambe risulta che Q-Q_1 e Q-Q_2 sono di uguale estensione e quindi equiscomponibili. Pertanto Q_1 e Q_2 sono equicompletabili. □

Osserviamo che le costruzioni effettuate nella dimostrazione del teorema sono tutte effettuabili con riga e compasso. Ad esempio nel dire che un triangolo è equiscomponibile ad un rettangolo, il rettangolo viene costruito considerando i punti medi di due lati. D'altra parte i punti medi di un segmento possono, appunto, essere trovati con il compasso. Questo significa che tali costruzioni rientrano nella nozione di equiscomponibilità anche se l'espressione "esiste una partizione" viene rafforzata col chiedere che i punti dove "tagliare" una figura siano tutti costruibili con riga e compasso.

Problema. Dato un foglio di carta, tagliarlo in modo da poter ricostruire un esagono regolare. Con lo stesso foglio ricostruire un pentagono.

Problema. Dato un triangolo, tagliarlo in modo da poter ricostruire un pentagono.

14. Equiscomponibilità e quadratura

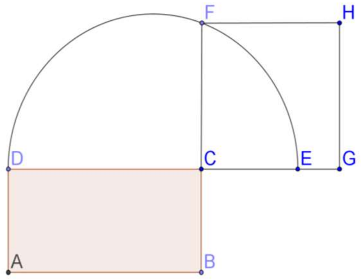
Abbiamo già detto che mentre attualmente "misurare" una cosa significa attribuire un numero reale a tale cosa per i matematici greci significava trovare un rapporto con altre cose ritenute più semplici. Tuttavia il teorema 12.4 comporta il seguente corollario che trasforma l'equiscomponibilità in una sorta di misura (una volta fissato un segmento unitario).

Proposizione 14.1. Si fissi un segmento come unità di misura, allora l'area di una figura poligonale $\underline{F}$ può essere calcolata scomponendo $\underline{F}$ e ricostruendo un rettangolo con base unitaria. L'altezza di tale rettangolo è la misura di $\underline{F}$.

In realtà, possiamo anche fare a meno di fissare una unità di misura se ci riferiamo ai quadrati invece che ai rettangoli. Questo perché mentre di rettangoli di una data estensione ve ne sono infiniti, di quadrati ne esiste uno ed uno solo o, per meglio dire, ne esistono infiniti ma tutti uguali tra loro.

Proposizione 14.2. Ogni rettangolo è equiscomponibile ad un quadrato. Pertanto ogni figura poligonale $\underline{F}$ è equiscomponibile ad un quadrato.

Dim. Indichiamo con a e b le lunghezze dei lati del rettangolo, il quadrato che vogliamo ottenere deve avere un lato la cui lun-



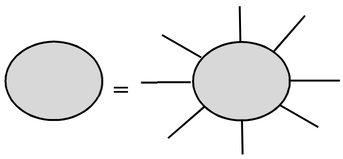
ghezza l deve essere tale che $l^2 = a \cdot b$. In altre parole il lato deve essere medio proporzionale tra DC e CB. Questo suggerisce di utilizzare il secondo teorema di Euclide che afferma appunto che l'altezza di un triangolo rettangolo è medio proporzionale tra le proiezioni dei cateti sull'ipotenusa (nel secondo capitolo se ne riporta la dimostrazione). Prolunghiamo allora il lato DC con un segmento CE uguale a CB e tracciamo la semicirconferenza di diametro DE. Detto F il suo punto di incontro con la retta CB, l'altezza CF del triangolo DFE è medio proporzionale tra DC e CE. Questo significa che il quadrato $CGHF$ ha la stessa estensione di $ABCD$ e che è quindi equiscomponibile con $ABCD$. □

Ovviamente il quadrato a cui si perviene equiscomponendo $\underline{F}$ è 'unico' nel senso che due quadrati equiscomponibili alla stessa figura sono uguali tra loro. Questo fatto permette di interpretare la tecnica dell'equiscomponibilità come un modo per misurare le figure geometriche a contorni rettilinei. Per fare questo bisogna vedere l'insieme dei quadrati come struttura algebrica, cosa non difficile. Basta definire la somma di due quadrati Q_1 e Q_2 spostando i quadrati in modo da formare due cateti di un triangolo rettangolo e poi porre Q_1+Q_2 uguale al quadrato Q costruito sull'ipotenusa di tale quadrato. Inutile dire che Q, avendo la stessa estensione della figura ottenuta unendo Q_1 con Q_2 è equiscomponibile con tale unione. In realtà per costruire correttamente la struttura algebrica dovremmo essere molto più precisi. Infatti si dovrebbe specificare che non si lavora con figure geometriche ma con figure geometriche definite a meno della nozione di congruenza e quindi con classi complete di un opportuno quoziente. Inoltre nelle due classi che si intendono sommare devono essere scelti due quadrati disgiunti tra loro. Tuttavia non mi soffermo su tali particolari. E' evidente che la struttura che si ottiene è isomorfa alla struttura additiva dei reali positivi.

15. Equiscomponibilità ed intuizione

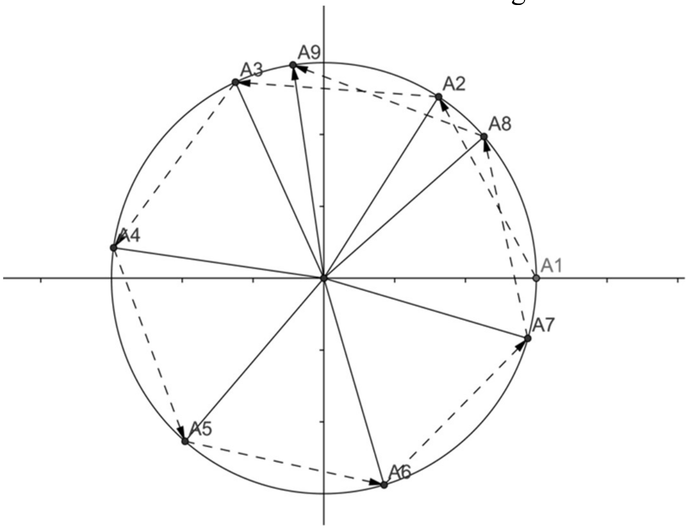
Il metodo delle equiscomponibilità per il calcolo delle aree è notevolmente intuitivo. Tuttavia non sempre è corretto fidarsi dell'intuizione come mostra la seguente proposizione.

Proposizione 15.1. (**Paradosso degli otto raggi**). Un cerchio di raggio r è equiscomponibile ad un cerchio unito ad otto raggi di lunghezza r.



Dim. Supponiamo che il cerchio C abbia centro nell'origine O e raggio 1 e sia $A_1, A_2, \ldots, A_n$ ... la successione di punti tale che A_1 sia il punto di coordinate (1,0) ed tale che A_i si ottiene ruotando A_{i-1} in senso antiorario di un angolo la cui misura in radianti sia 1. In altre parole, A_n è il punto della circonferenza tale che OA_n forma un angolo di n-1 radianti con l'asse delle ascisse. I punti di tale successione sono tutti diversi tra loro poiché se fosse $A_n = A_m$ dovrebbe essere n-m un multiplo di 2π. Ma ciò è assurdo poiché da una equazione del tipo $n\text{-}m = 2\pi q$ seguirebbe che $\pi = (n\text{-}m)/2q$ in contrasto con l'irrazionalità di π. Indichiamo con OA_n l'insieme dei punti interni al segmento che congiunge O con A_n e chiamiamo *raggiera* l'insieme $Ra = \bigcup_{n\in N} A_n$. Consideriamo ora la figura C-Ra che si ottiene togliendo dal cerchio la raggiera Ra e spostiamola ottenendo un cerchio C^* "con tagli".

Togliamo da *Ra* i primi 8 raggi (che mettiamo da parte) e ruotiamo la figura ottenuta di 8 radianti in senso orario. Otteniamo di nuovo *Ra* che opportunamente spostata possiamo aggiungere a C^*. In tale modo si ottiene un cerchio C^{**} uguale al cerchio C



di partenza. Tuttavia in tale operazione ci sono rimasti ancora 8 raggi che possiamo liberamente spostare verso C^* per ottenere la forma che abbiamo indicato in figura.[23] □

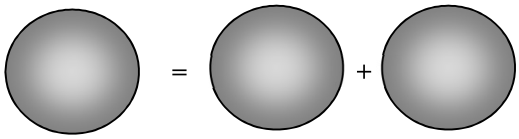
Da notare che tale proposizione è compatibile con il fatto che figure equiscomponibili hanno uguale estensione. Infatti gli otto segmenti hanno misura nulla. Nel capitolo 2 nel paragrafo sui falsi teoremi mostreremo un ulteriore esempio di come l'intuizione sulla equiscomponibilità conduca ad errore.

Da notare che un esempio clamoroso di equiscomponibilità completamente contrario al senso comune è il seguente. Esso è noto sotto il nome di "paradosso di Banach-Tarski" dal nome dei suoi autori.

Teorema 15.2. (Paradosso di Banach-Tarski). Se si accetta l'assioma della scelta, una sfera di raggio unitario è equiscomponibile ad una coppia di sfere di raggio unitario.

---

[23] La struttura di tale dimostrazione non è molto diversa da quella utilizzata nel "gioco" degli alberghi di Hilbert esposta alla fine del capitolo 1 del secondo volume (si veda la relativa nota).



Tale teorema viene anche chiamato “paradosso dei pani e dei pesci” poiché ricorda il famoso miracolo raccontato nei vangeli, miracolo che sembrerebbe quindi logicamente possibile. Non provo ad esporre la dimostrazione che è bella ma troppo lunga. Osservo solo che se le n parti in cui viene spezzata la prima sfera avessero volume $v_1, \dots, v_n$, allora $1 = v_1 + \dots + v_n = 2$.

16. Contro i matematici

Quando ho detto che la matematica giocava un ruolo centrale nella cultura dei greci non intendevo dire che ciò accadeva per tutti i greci. In realtà, essendo gli antichi greci un popolo notevolmente intelligente, vivace e non conformista, accadeva che su di un dato argomento ciascuno avesse una propria idea. Così sulla matematica non tutte le opinioni erano concordi e lo stesso Aristotele non era convinto della importanza della matematica allo stesso modo di Platone.

Uno di quelli che meno avevano in simpatia la matematica era Sesto Empirico, un filosofo vissuto attorno al 180 dopo Cristo. Nel suo libro "Contro i geometri" egli critica in maniera radicale l'opera dei matematici. Ad esempio critica le basi logiche del metodo ipotetico deduttivo di Euclide e precisamente la validità di un tale metodo come strumento per ottenere verità sul mondo.

Infatti per Sesto Empirico il valore conoscitivo di un sistema ipotetico-deduttivo non può stare nelle ipotesi (cioè nei postulati) in quanto o una asserzione è vera, ed allora non ha senso considerarla come ipotesi, oppure è falsa ed allora è sbagliato prenderla come ipotesi.

> *“. . . la cosa ammessa per ipotesi o è vera e tale come noi la supponiamo, o è falsa. Ma se essa è vera, noi non la postuliamo, perché non sentiamo il bisogno di ricorrere ad una cosa piena di sospetto quale è l'ipotesi, ma l'assumiamo immediatamente, giacché nessuno assume ipoteticamente le cose vere ed esistenti, quali ad esempio il fatto che adesso è giorno ed io sto*

discutendo e respirando . . . Se però essa non è tale, ma è falsa, non si ricava alcun vantaggio dell'ipotesi . . . "

Si potrebbe allora dire che il valore di un sistema ipotetico deduttivo stia nella validità delle conseguenze che se ne ricavano.

"Ma, per Zeus, essi (i matematici) dicono, se quello che segue dall'ipotesi si scopre essere vero, senz'altro saranno vere anche le cose assunte in via ipotetica, ossia le cose da cui quelle vere seguono".

Questo è il punto di vista della fisica moderna; nessun fisico pensa che le leggi generali di una teoria fisica siano direttamente verificabili o applicabili. Piuttosto si considera come una conferma della validità di una teoria fisica il fatto che le conseguenze di tale teoria siano state verificate. Ad esempio, se ci si riferisce ai principi della dinamica classica di Newton, allora una prova di tali principi è stata vista nella loro capacità di prevedere il moto dei pianeti non nel fatto che siano stati effettuati esperimenti di laboratorio capaci di verificarli. In breve, nella fisica si procede stabilendo una serie di assunzioni (che possiamo chiamare leggi, ipotesi, assunzioni, postulati, assiomi) che in generale non sono di per se stesse verificabili o utili, poi da queste assunzioni si ricavano altre assunzioni (i teoremi) che possono essere assoggettate ad una verifica. Se tali teoremi dopo la verifica risultano veri allora si dirà che la teoria è valida. La struttura di tale modo di procedere è la seguente

da B e $A \Rightarrow B$ segue A;

e non deve essere confusa con la regola di Modus Ponens che invece ha la struttura

da A e $A \Rightarrow B$ segue B;.

Ma ciò che viene considerato soddisfacente dai moderni scienziati non soddisfa invece Sesto che, giustamente, osserva che dal fatto che da alcune ipotesi si siano tratte conseguenze vere non si può dedurre che tali ipotesi siano vere.

"Ebbene, anche una tale affermazione risulta ancora semplicistica . . . se il conseguente è vero, non per questo è tale anche il precedente . . . come al fatto che la terra vola (*il ché è falso*) *consegue che la terra esiste* (*il ché è vero*)*."*

A parte l'infelice scelta dell'esempio (sappiamo adesso che la terra vola), Sesto aveva ragione. Ad esempio dalla equazione 2=5 moltiplicando entrambi i membri per zero si ricava che $0\times2=0\times5$ cioè che 0 = 0. Pertanto abbiamo un esempio in cui l'implicazione

$$2=5 \Rightarrow 0 = 0$$

è vera, la conseguenza 0 = 0 è vera ma l'ipotesi 2 = 5 è falsa. D'altra parte vi sono moltissimi esempi di teorie che con il tempo si sono dimostrate false ma che hanno prodotto teoremi veri. Un caso è dato dalla teoria degli insiemi che ha prodotto molti ed utili teoremi nonostante che la scoperta dei paradossi abbia dimostrato la sua falsità. D'altra parte ogni teoria fisica è stata dimostrata essere falsa dalla teoria successiva, in un certo senso.
Altre obiezioni sono inerenti direttamente alla matematica e precisamente alla concezione degli enti ideali.

> *" . . . essi dicono, inoltre, che una linea viene prodotta dallo scorrimento di un punto, una superficie dallo scorrimento di una linea, e un corpo solido dallo scorrimento di una superficie . . . "*

> *" . . . il punto che essi definiscono come segno-privo-di-dimensioni, si deve concepire o come corporeo o come incorporeo. Corpo esso non è, secondo le loro stesse affermazioni, giacché le cose che non hanno dimensione, secondo loro, non sono corpi. Resta allora da dire che esso è incorporeo, il che è ancora una vola incredibile. Infatti ciò che è incorporeo non si può concepire come generatore di una linea; quindi il punto non è un segno-privo-di-dimensioni."*

In altri termini Sesto si pone il problema di come un ente senza dimensioni, il punto, possa generare (per scorrimento) un ente con una dimensione, la linea. Tale osservazione equivale, in un certo senso, ad osservare che se la lunghezza di un punto è zero allora ogni segmento, in quanto insieme (somma) di punti, deve avere lunghezza zero. La differenza consiste nel riferirsi alla linea prodotta dallo scorrimento di un punto (operazione questa che non sembra coinvolgere l'infinito attuale) e non alla linea intesa come insieme di punti (concezione questa che, coinvolgendo l'infinito attuale, non era presa in considerazione). Tale tipo di argomentazione viene ripetuta anche per confutare il concetto di linea senza larghezza il cui scorrimento genera una superficie e quello di superficie senza spessore il cui scorrimento genera i solidi.

Infine un altro tipo di critica riguarda il procedimento di astrazione mediante il quale l'uomo perverrebbe a concepire gli enti matematici ideali. Infatti per Sesto tutto ciò che viene concepito viene concepito o mediante una diretta esperienza oppure tramite un procedimento di immaginazione-astrazione. Ora è evidente che non possiamo mai avere una esperienza diretta, ad esempio, di una linea senza larghezza e che questa idea di linea non è simile a niente di esistente

> *. . . giacché non cade sotto i nostri sensi una lunghezza che sia priva di larghezza . . .*

D'altra parte un procedimento di immaginazione-astrazione può avvenire

> *. . . per somiglianza, ad esempio dall'immagine di Socrate lo stesso Socrate, per composizione, ad esempio dal cavallo e dall'uomo un ippocentauro giacché mescolando le membra del cavallo e dell'uomo noi siamo riusciti ad immaginare l'ippocentauro che non è né uomo né cavallo ma è composto da entrambi. Per analogia, infine, si concepisce qualcosa ancora in due guise ossia o per accrescimento o per diminuzione, come quando, ad esempio, tenendo presenti gli uomini normali . . . concepiamo per accrescimento il Ciclope . . . e come quando per diminuzione immaginiamo un pigmeo che non ci è mai caduto sotto i sensi.*

Ora è evidente che il concetto di linea senza larghezza non è simile a niente di esistente per lo stesso motivo per cui non è determinato dall'esperienza. E' anche immediato che tale concetto non si ottiene per composizione come nel caso dell' ippocentauro. Non resta altro che il procedimento di diminuzione ma anche questo permette solo di ridurre a piacere (potenzialmente) la larghezza non di considerarla (attualmente) nulla e quindi non permette il tipo di astrazione che si richiederebbe. Anche in questo caso entra in gioco il rifiuto dell'infinito attuale. Ad esempio se noi volessimo definire un punto P immaginando una piccola sfera s_1 di centro P, poi un'altra sfera s_2 di centro P e raggio dimezzato e così via, allora il punto P sarebbe il frutto del processo di astrazione definito in tale modo solo se noi potessimo considerare tale processo terminato con un operazione di limite che sarebbe possibile solo se si accettasse l'infinito attuale.

Un altro tipo ancora di procedimento possibile di astrazione è quello che fa pervenire ad un concetto mediante una semplice cancellazione di alcune delle proprietà dell'oggetto di partenza. Così il concetto di linea senza larghezza si può ottenere semplicemente "facendo finta" che la larghezza di un oggetto reale non esiste. Ma:

> *. . . se noi, dopo aver concepito una certa lunghezza avente una data quantità di larghezza, abbiamo altresì la possibilità di assumere una lunghezza priva di larghezza sopprimendo quest'ultima, allora allo stesso modo, dopo aver concepito un pezzo di carne che abbia la proprietà di essere vulnerabile, mediante la soppressione di tale proprietà noi potremo anche concepire una carne che non sia soggetta alla vulnerabilità . . . Ma tale cosa è completamente impossibile e contraria alle comuni nozioni umane: infatti ciò che viene concepito come invulnerabile secondo noi non è affatto carne, giacché la carne, in quanto carne, viene concepita con la proprietà di essere vulnerabile . . . Onde anche la lunghezza concepita come priva di larghezza non potrebbe essere una lunghezza, giacché la lunghezza, in quanto lunghezza, viene concepita come avente una certa quantità di larghezza.*

In altre parole se un oggetto concreto A venga rappresentato da un oggetto ideale $\underline{A}$ facendo astrazione da alcune particolari caratteristiche di A allora non è chiaro che relazione sussiste tra A ed $\underline{A}$. In particolare, supponiamo di avere dimostrato che una proprietà vale per $\underline{A}$, allora chi ci autorizza a dire che tale proposizione è valida anche per A? Ad esempio supponiamo che un astronomo debba studiare l'orbita di un asteroide A in presenza dei pianeti $A_1, A_2,..., A_n$ e che per fare questo decida di rappresentare con il punto $\underline{A}$ l'asteroide e con i punti $\underline{A}_1,..., \underline{A}_n$ i pianeti (su cui si concentra la massa). Immaginiamo inoltre che, applicando la geometria euclidea e la meccanica del moto dei punti, giunga ad una qualche conclusione P relativa a $\underline{Z}$, chi ci assicura che P sia valida anche per l'asteroide reale A?

LETTURA

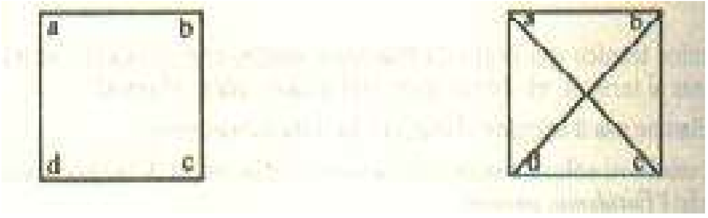
Platone e la duplicazione del quadrato (da *IL MENONE*).[24]

MENONE: Va bene, caro Socrate, ma in che senso tu sostieni che noi non apprendiamo ma che ciò che noi chiamiamo “apprendimento” è reminiscenza? Sapresti insegnarmi che è veramente così?
SOCRATE: Già prima dicevo, caro Menone, che sei un furbacchione, ora mi domandi se so insegnarti proprio mentre sto dicendo che non c'è insegnamento ma reminiscenza evidentemente per farmi subito apparire in contraddizione con me stesso.
MENONE: No, per Zeus, o Socrate, non l'ho detto con questo scopo, ma solo per l'abitudine. Se, però, in qualche modo mi puoi dimostrare che la cosa sta così come dici, allora dimostramelo.
SOCRATE: Non è facile! Tuttavia, per te, sono disposto a farlo. Chiamami un po' uno dei tuoi numerosi servi che sono qui, quello che vuoi tu, perché tramite lui ti possa dare la dimostrazione.
MENONE: Certo. Vieni qui, ragazzo!
SOCRATE: E' greco e parla greco?
MENONE: Si, perfettamente. E' nato in casa.
SOCRATE: Fa' bene attenzione, se ti sembra che si ricordi o che impari da me.
MENONE: Presterò attenzione.
SOCRATE: Dimmi un po', ragazzo, sai che questa qui è un'area quadrata (abcd)?
RAGAZZO: Si.

---

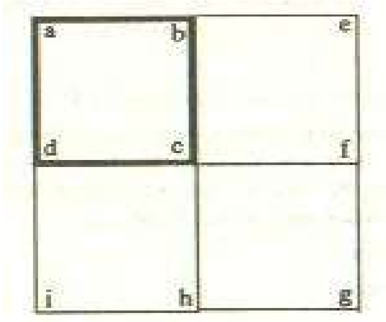
[24] In questo dialogo si dimostra un metodo per duplicare un quadrato. La dimostrazione equivale, in un certo senso, alla dimostrazione del teorema di Pitagora nel caso particolare di un triangolo rettangolo i cui cateti sono uguali. Infatti se i cateti misurano c e l'ipotenusa misura i, allora il fatto che $2 \cdot c^2 = i^2$ equivale a dire che il quadrato costruito sull'ipotenusa è il doppio del quadrato costruito su un cateto. Da notare che la duplicazione del cubo è un problema notevolmente più complesso ed è stato dimostrato che non può essere affrontato con i soli strumenti della riga e compasso. Tale dialogo, in cui si vuole giustificare la teoria della reminiscenza, costituisce un bellissimo esempio di “didattica della matematica” in quanto lo scopo di Socrate è quello di fare “emergere” la dimostrazione del teorema dalla mente del servo.

SOCRATE: Il quadrato è dunque una superficie che ha uguali tutti questi lati, che sono quattro (ab, bc, cd, da).

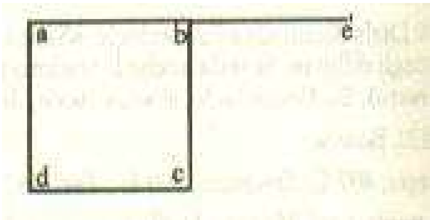


RAGAZZO: Certamente.
SOCRATE: E non ha forse uguali anche queste linee qui, che lo attraversano nel mezzo (ac, bd)?
RAGAZZO: Sì.
SOCRATE: E non potrebbe esserci forse una superficie come questa e più grande e più piccola?
RAGAZZO: Certamente.
SOCRATE: Se dunque questo lato (bc) fosse di due piedi, e anche questo (ab) di due, di quanti piedi sarebbe l'intero? Fa questa considerazione: se da questa parte (ab) fosse di due piedi e da quest'altra (bc) di uno solo, la superficie non sarebbe forse di una volta due piedi?
RAGAZZO: Sì .
SOCRATE: Ma, poiché anche da questa parte (bc) è di due piedi, non diventa di due volte due piedi?
RAGAZZO: Sì, diventa.
SOCRATE: Diventa, perciò, di due volte due piedi?
RAGAZZO: Esatto .
SOCRATE: E quanti sono, allora, due volte due piedi? Fa' il conto e dillo.
RAGAZZO: Quattro, o Socrate.

SOCRATE: E non potrebbe darsi un'altra superficie doppia di questa, ma tale da avere tutti i lati eguali come questa?



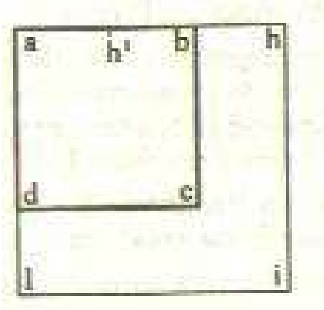
RAGAZZO: Sì.
SOCRATE: Di quanti piedi sarà dunque?
RAGAZZO: Di otto.
SOCRATE: E ora cerca di dirmi di quanto sarà ciascun lato di essa. Il lato di questa è di due piedi; e, allora, di quanto sarà quello di quella doppia?
RAGAZZO: E' chiaro, o Socrate, che sarà doppio.
SOCRATE: Vedi, caro Menone, che io non gli insegno, ma che lo interrogo su ogni cosa? Ed ora, costui ritiene di sapere quale sia il lato dal quale deriverà l'area di otto piedi: o non ti sembra?
MENONE: A me sì.
SOCRATE: E lo sa, dunque?
MENONE: Per nulla.
SOCRATE: Però ritiene che derivi dal lato doppio.
MENONE: Sì.
SOCRATE: Osserva come verrà via via ricordandosi, come appunto deve ricordarsi. E tu dimmi: dal lato doppio, dici che ha origine la superficie doppia? E tale, dico, che non sia di qui lunga e di qui corta, ma che sia eguale da ogni parte come questa qui, però doppia di questa, ossia di otto piedi. Ma sta' attento, se ti sembra ancora che possa derivare dal lato doppio.



RAGAZZO: A me sì.
SOCRATE: E non diventa forse questo lato (ae) doppio di questo (ab), se ne aggiungiamo un altro come questo, da questa parte (be)?
RAGAZZO: Certamente.
SOCRATE: Da questo (ae), dici tu, deriverà la superficie di otto piedi, quando si tracceranno quattro lati come questi .

RAGAZZO: Esattamente.
SOCRATE: Ma in questa superficie non ci sono forse queste quattro qui (abcd, befc, cfgh, dchi), delle quali ognuna è uguale a questa di quattro piedi (abcd)?
RAGAZZO: Sì.
SOCRATE: Disegniamo, allora, a partire da questo 4 lati uguali . E' oppure no questa la superficie (aegi) che tu affermi essere di 8 piedi ?
RAGAZZO : Esattamente .
SOCRATE : Ma in questa superficie non vi sono forse queste 4 qui (abcd , befc , cfgh , dchi), delle quali ognuna é uguale a questa di 4 piedi?
RAGAZZO : Sì
SOCRATE: E quanto diventa allora? Non diventa quattro volte questa?
RAGAZZO: E come no?
SOCRATE: E allora, è il doppio quattro volte tanto?
RAGAZZO: No, per Zeus.
SOCRATE: Ma quante volte?
RAGAZZO: Quadruplo.
SOCRATE: Dunque, dal lato doppio, o ragazzo, non deriva una superficie doppia ma quadrupla.
RAGAZZO: Dici il vero.
SOCRATE: E quattro volte quattro, fanno sedici, no?
RAGAZZO: Sì.
SOCRATE: E allora, quella di otto piedi da quale lato? Non se ne ottiene da questo (ae) una quadrupla?
RAGAZZO: Sì, lo dico.
SOCRATE: E quella di quattro, dalla metà di questo qui (ae)?
RAGAZZO: Sì.
SOCRATE: Ebbene, l'area di otto piedi non è forse doppia di questa qui (abcd), e metà di quest'altra (aegi)?
RAGAZZO: Sì.
SOCRATE: E allora, non deriverà da un lato maggiore rispetto a questo (ab), ma minore rispetto a quest'altro (ae); o no?
RAGAZZO: Così mi pare.
SOCRATE: Bene; quello che a te sembra devi rispondere. E dimmi: questo lato (ab) non era di due piedi e quest'altro (ae) di quattro?
RAGAZZO: Si.

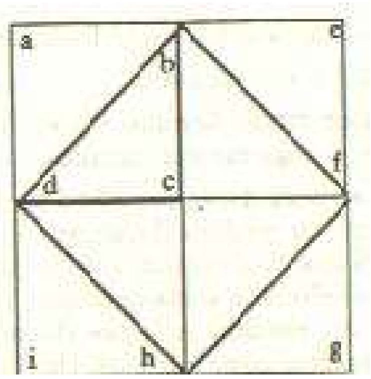
SOCRATE: Bisogna allora che il lato della superficie di otto piedi sia maggiore di questo di due, ma minore di quello di quattro.
RAGAZZO: Bisogna .
SOCRATE: Cerca allora di dire di che lunghezza tu affermi che esso debba essere.
RAGAZZO: Di tre piedi.
SOCRATE: Se deve essere di tre piedi, aggiungiamo dunque a questo lato (ab) la metà di questo (ah), e avremo i tre piedi (ah). Questi sono due piedi (ah) e questo uno (hh). Alla stessa maniera, a partire di qua si ottengono due piedi (ab) più un piede (dc). Ne deriva, così, l'area che tu dici (abil).
RAGAZZO: Sì.



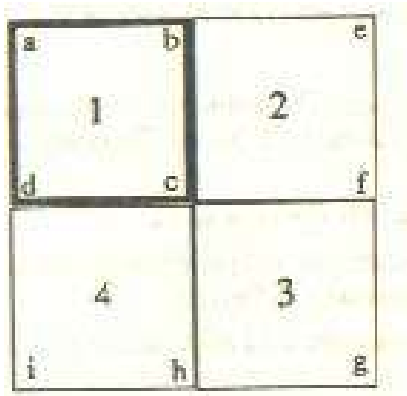
SOCRATE: Ma se da questa parte (ab) è di tre, e da quest'altra (hi) di tre, l'intera superficie non diventa di tre volte tre piedi?
RAGAZZO: Sembra.
SOCRATE: E tre volte tre, quante volte sono?
RAGAZZO: Nove.
SOCRATE: E il doppio, di quanti piedi doveva essere?
RAGAZZO: Otto.
SOCRATE: Dal lato di tre piedi non deriva per nulla la superficie di otto.
RAGAZZO: Certamente no.
SOCRATE: Ma allora da quale lato? Cerca di dircelo con esattezza; e se non vuoi fare calcoli, indicaci almeno da quale.
RAGAZZO: Ma per Zeus, o Socrate, io non lo so.
SOCRATE: Comprendi ora, o Menone, a che punto si trova attualmente nel processo del ricordare? Prima, cioè, non sapeva quale fosse il lato del quadrato di otto piedi, come del resto neppure ora lo sa; tuttavia, allora credeva di saperlo, e rispondeva con sicurezza come se sapesse e non riteneva di aver dubbi; ora è convinto di aver dubbi e come non sa, così neppure crede di sapere.
MENONE: Dici il vero.
SOCRATE: Non si trova dunque, ora, in una situazione migliore, relativamente alla cosa che non sapeva?
MENONE: Anche questo mi pare:

SOCRATE: Avendolo fatto dubitare, pertanto, e avendolo fatto intorpidire come fa la torpedine, gli abbiamo forse nuociuto?
MENONE: Non mi pare.
SOCRATE: Dunque, come sembra, gli abbiamo recato giovamento, al fine della ricerca di come stia effettivamente la cosa. Ora, infatti, ricercherebbe anche di buon grado, dal momento che non sa; mentre allora, facilmente, di fronte a molti e spesso avrebbe creduto di dire bene, affermando che per ottenere una superficie doppia, bisogna prendere il lato doppio in lunghezza.
MENONE: Sembra.
SOCRATE: Credi, dunque, che egli si sarebbe messo a cercare o ad imparare ciò che egli riteneva di sapere non sapendolo, prima che fosse caduto nel dubbio ritenendo di non sapere, e che avesse desiderato di conoscere?
MENONE: Non mi pare, o Socrate.
SOCRATE: Dunque, l'intorpidimento gli ha giovato?
MENONE: Mi sembra.
SOCRATE: Osserva, ora, da questo dubbio come scoprirà la verità, ricercando insieme a me, mentre io non farò altro che interrogarlo, senza insegnargli. . E fa bene attenzione che tu non mi colga ad insegnargli o a spiegargli, e non solo ad interrogarlo intorno alle sue convinzioni.
Dimmi, dunque: non è di quattro piedi questa superficie (abcd)? Comprendi?
RAGAZZO: Sì.
SOCRATE: Potremmo aggiungere ad essa quest'altra eguale (befc)?
RAGAZZO: Sì.
SOCRATE: E quest'altra terza, uguale a ciascuna di queste (cfgh)?
RAGAZZO: Sì.
SOCRATE: E non potremmo anche completare la figura in questo angolo (dchi)?

RAGAZZO: Certamente.
SOCRATE: E non risulteranno queste quattro superfici eguali ?
RAGAZZO: Sì.
SOCRATE: E, allora, tutto questo intero (aegi), quante volte diventa più grande di questo (abcd)
RAGAZZO: Quattro volte.
SOCRATE: Per noi, invece, doveva essere il doppio; o non ricordi?



RAGAZZO: Certamente.
SOCRATE: E questa linea tracciata da un angolo all'altro (bd, bf, fh, hd), non viene forse a dividere a metà ciascuna di queste superfici?
RAGAZZO: Sì.
SOCRATE: Non si ottengono, dunque, queste quattro linee uguali racchiudenti quest'area qui (bfhd)?
RAGAZZO: Sì, si ottengono.
SOCRATE: Considera allora: quanto grande è questa superficie (bfhd)?
RAGAZZO: Non lo so.
SOCRATE: Di questi quadrati, che sono quattro, ciascuna linea non ha tagliato internamente la metà di ciascuno? O no?
RAGAZZO: Sì.
SOCRATE: E quante ve ne sono di queste metà in questa figura (bfhd)?
RAGAZZO: Quattro.
SOCRATE: E quante in quest'altra (abcd)?



RAGAZZO: Due.
SOCRATE: E il quattro che cos'è rispetto al due?
RAGAZZO: Il doppio.
SOCRATE: Questa superficie, dunque, di quanti piedi diventa?
RAGAZZO: Di otto piedi.

SOCRATE: Da quale linea?
RAGAZZO: Da questa (a'b).
SOCRATE: Da quella che abbiamo tracciata da un angolo all'altro del quadrato di otto piedi?
RAGAZZO: Sì.
SOCRATE: Coloro che se ne intendono chiamano questa linea diagonale; sicché, se essa ha nome diagonale, allora dalla diagonale, come tu dici, o ragazzo di Menone, si può ottenere l'area doppia.
RAGAZZO: Certamente, o Socrate.
SOCRATE: Che cosa ti sembra, o Menone? C'è qualche pensiero da lui espresso che non sia suo ?
MENONE: No, tutti suoi.
SOCRATE: Eppure, non sapeva, come dicevamo poco fa.
MENONE: Dici il vero.
SOCRATE: E c'erano in lui questi pensieri o no?
MENONE: Sì.
SOCRATE: Dunque, in chi non sa intorno alle cose che non sa, vi sono opinioni vere che ad esse si riferiscono?
MENONE: Sembra.
SOCRATE: Ora in lui, come un sogno, sono state suscitate queste opinioni; e, interrogandolo di nuovo più volte e in molti modi su queste stesse cose, sta certo che finirà per sapere con precisione, sulle medesime, non meno esattamente di ogni altro .
MENONE: Pare proprio di sì.
SOCRATE: Dunque, egli saprà senza che nessuno gli insegni, ma solo che lo interroghi, traendo egli stesso la scienza da se medesimo.
RAGAZZO: Sì.
SOCRATE: E questo trarre la scienza di dentro a sé, non è ricordare?
MENONE: Certamente.
SOCRATE: E la scienza che ora egli possiede, o la imparò un tempo o la possedette sempre.
MENONE: Sì.
SOCRATE: Dunque, se la possedette sempre, fu anche sempre conoscente; e se, invece, l'ha appresa in un tempo, non poté certo averla appresa nella presente vita. Oppure gli insegnò qualcuno geometria? Costui, infatti, farà lo stesso per tutta la geometria, e

per tutte quante le altre scienze. C'è, forse, uno che gli abbia insegnato tutto? A buon diritto tu devi saperlo: non per altro, perché è nato ed è stato allevato in casa tua.
MENONE: Ma lo so che nessuno gli ha mai fornito insegnamenti.
SOCRATE: Ed ha o non ha queste conoscenze?
MENONE: Necessariamente, o Socrate, sembra.
SOCRATE: Allora, se non le ha acquisite nella presente vita, questo non è ormai evidente, ossia che le ebbe e le apprese in un altro tempo?
MENONE: E' chiaro.
SOCRATE: E non è forse questo il tempo in cui egli non era uomo?
MENONE: Sì .
SOCRATE: Se, allora, e nel tempo in cui è uomo e nel tempo in cui non lo è, vi sono in lui opinioni vere, le quali, risvegliate mediante l'interrogazione, diventano conoscenze, l'anima di lui non sarà stata in possesso del sapere sempre in ogni tempo? E' evidente infatti che, nel corso di tutto quanto il tempo, talora è e talora non è uomo.
MENONE: E' chiaro.
SOCRATE: Se, dunque, sempre la verità degli esseri è nella nostra anima, l'anima dovrà essere immortale. Sicché bisogna mettersi con fiducia a ricercare ed a ricordare ciò che attualmente non si sa: questo è infatti ciò che non si ricorda.
MENONE: Mi sembra che tu dica bene, o Socrate, ma non so come.

LETTURA

Riporto il testo di un mio articolo divulgativo
G. Gerla, Un punto dal volto di gatto, *Periodico di Matematiche*, 3,4, (2006) 9-20, 15-25. [25]

1. Introduzione

There are more points in Heaven and Heart, Horatio, than are dreamt of in our geometry (pseudo-Shakespeare, **Amleto**).

Punto di partenza di questo articolo è la domanda se sia possibile definire una teoria matematica in cui abbia senso una affermazione del tipo:
"un punto dall'aspetto di gatto ruotò verso me sorridendo".
Più in generale, ci chiediamo:

è possibile allargare l'universo dei punti, definendo punti che assomigliano ad un gatto, punti triangolari, punti con direzione, punti che ruotano su se stessi ?

Sembra difficile poter dare una risposta positiva a tale questione. Infatti, come è noto, nella geometria che abbiamo appreso a scuola e all'università tutti i punti sono uguali tra loro e quindi non può esserci un punto dall'aspetto di gatto ed un punto, ad esempio, dall'aspetto di cane. Detto in termini più formali, il gruppo delle isometrie è transitivo cioè dati due punti P e Q esiste una isometria che porta P in Q. Ciò comporta che tutto quello che si può dire di P si può dire anche di Q e viceversa. Inoltre i punti non possono girare su se stessi come invece possono fare le figure solide perché ad un punto non si sa come "attaccare" un sistema di riferimento. D'altra parte, se il punto gira su se stesso come fare per distinguere la posizione del punto prima e dopo la rotazione ? Insomma: un pasticcio.

[25] In tale articolo si pone il problema se la scelta di Euclide e di tutti i matematici successivi di considerare il punto come ente primitivo sia una scelta necessaria oppure no.

Naturalmente se avessi scritto:

"un solido con la forma di gatto che sorride ruotò verso me"

vi sarebbero stati (forse) meno problemi e l'usuale geometria fornitaci da Euclide sarebbe stata (forse) sufficiente a dare un significato a tale frase. Ma anche in questo caso le cose non sono così semplici come appaiono in quanto non è affatto chiaro che cosa si debba intendere, da un punto di vista matematico, per "corpo solido". A dispetto di quanto si pensa comunemente, la geometria non dice molto in proposito poiché si occupa di regioni dello spazio e non di corpi solidi, di posizioni nello spazio e non di punti materiali. Infatti, qualunque idea si abbia di che cosa sia un corpo, è certo che un corpo è qualcosa di invariante rispetto ai possibili movimenti nello spazio ed è comunque diverso dalla regione dello spazio che esso può occupare in un dato istante. Insomma anche la questione di che cosa sia un corpo non è banale (si vedano ad esempio [14], [3], [4]).

Ma torniamo ai punti. Come abbiamo già osservato, le formalizzazioni proposte nelle teorie geometriche usuali non sono l'ambito adatto per andare a cercare punti dalla forma di gatto. Questo significa che dobbiamo andare a cercare "ambienti matematici" in cui i punti possano essere definiti a nostro piacimento e non trovarli, in quanto nozioni primitive, già belli e fatti. Tali ambienti dovrebbero permettere di capovolgere il punto di vista della geometria usuale in cui, a partire dai punti, si definiscono regioni di tipo sferico, cubico e così via. Si dovrebbe invece partire (ad esempio) dalle regioni, e definire i punti (magari di tipo sferico, cubico e così via). In definitiva si pone la questione se il considerare i punti come enti primitivi costituisca una scelta necessaria per ogni trattazione rigorosa della geometria. Il fatto è che la nozione di punto è forse quella più misteriosa (ed importante) della matematica e ciò perché a tale nozione è legato uno dei processi fondamentali del pensiero razionale, il processo di astrazione.

2. Regioni, inclusione e contatto tra regioni

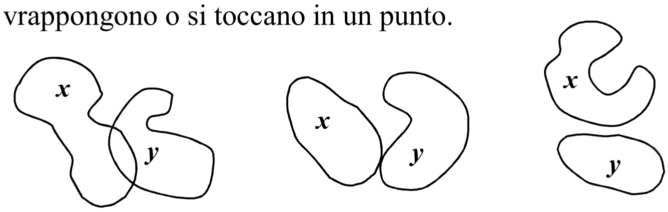
Il primo a fondare la geometria sul concetto di corpo solido fu probabilmente Lobačevskij il quale vedeva i punti, le linee e le superfici come tipi di *contatto* tra i corpi. Purtroppo le definizioni di Lobačevskij sono alquanto oscure e, pur essendo presente l'enunciazione per cui i punti sono definiti a partire dai solidi, di fatto l'approccio proposto continua a basarsi sulla nozione di punto. Più interessanti sono i tentativi del famoso matematico e filosofo A. N. Whitehead che molti anni dopo inizia un ampio esame di come si possa derivare la nozione di punto da quella di regione. Appaiono le opere, *An Inquiry Concerning the Principle of Natural Knowledge* (1919) e *The concept of Nature* (1920) in cui sono assunte come primitive le nozioni di "regione" e di "inclusione tra regioni". L'inclusione non viene intesa in senso insiemistico ma come una relazione d'ordine nell'insieme delle regioni soddisfacente particolari proprietà. D'altra parte la definizione insiemistica di inclusione ($X \subseteq Y$ se ogni elemento di X è un elemento di Y) non corrisponde affatto all'idea intuitiva che, ad esempio, permette di dire che il volante è un pezzo dell'auto o che un ritornello è un pezzo di una canzone.

Precisamente Whitehead considera strutture del tipo (M, K), dove M è un insieme i cui elementi chiama *eventi* e K una relazione binaria in M detta *relazione di estensione*. Tale relazione è assunta come primitiva e deriva dal concetto intuitivo "*contenere*". Per sottolineare l'interpretazione geometrica e non fisica del discorso di Whitehead preferisco sostituire il termine *evento* con quello di *regione* ed il termine *estensione* con quello duale di *inclusione* (non necessariamente propria) denotato con $\leq$. Possiamo tradurre le proprietà elencate da Whitehead nel seguente sistema di assiomi:

(*i*) $\leq$ è una relazione d'ordine
(*ii*) $\forall z \exists x \exists y (x<z<y)$
(*iii*) $\forall x \forall y (x<y \Rightarrow \exists z\ (x<z<y)$
(*iv*) $\forall x \forall y \exists z (x \leq z \wedge y \leq z)$
(*v*) $\forall x \forall y (\forall x' (x'<x \Rightarrow x'<y) \Rightarrow x \leq y)$.

Definizione 1. Si chiama *spazio di inclusione* ogni modello $(\mathfrak{R}, \leq)$ del sistema di assiomi *i*)-*v*).

La teoria che ne esce fuori risulta comunque alquanto insoddisfacente. Infatti esistono notevoli difficoltà nella definizione di punto quando ci si riferisca alla sola relazione di inclusione. Non è sorprendente quindi che dopo pochi anni in *Process and Reality* Whitehead sostituisca la nozione (di sapore insiemistico) di inclusione con quella (di sapore topologico) di connessione intesa come contatto o sovrapposizione. Come illustrato nelle prime due immagini della figura seguente, tale relazione deve essere vista come la relazione che sussiste tra due regioni che si sovrappongono o si toccano in un punto.



Nella figura sono riportate due coppie di regioni connesse tra loro ed una coppia che non è connessa. Da notare che le indagini di Whitehead, di carattere filosofico-epistemologico piuttosto che matematico, non sono e non vogliono essere una proposta di un sistema rigoroso di assiomi e definizioni. Esporrò tuttavia le idee basilari contenute in *Process and Reality* riferendomi alla formalizzazione proposta da Gerla e Tortora in [7].

Come dicevamo, alla base del discorso di Whitehead sono strutture del tipo $(\Re, C)$ con $\Re$ insieme i cui elementi vengono detti *regioni* e C relazione binaria in $\Re$ detta *relazione di connessione*. Se x è nella relazione C con y, dirò che *x è connesso ad y* e scriverò xCy. Inoltre indicherò con $C(x)$ l'insieme $\{y \in \Re : xCy\}$ delle regioni che sono connesse con x. Ciò consente di definire la relazione di inclusione tra regioni.

Definizione 2. Data una coppia di regioni x ed y, poniamo $x \leq y$ e diciamo che x è *contenuta* in y se $C(x) \subseteq C(y)$.

In altre parole, diciamo che x è contenuta in y se non è possibile entrare in contatto con x senza entrare in contatto anche con y. E' evidente che $\leq$ è un preordine, cioè che vale la proprietà riflessiva e transitiva. Non mi soffermerò su tutte le proprietà che Whitehead ritiene debbano essere verificate dalle strutture di connessione; si tenga presente che solo nel Capitolo 2 ne sono indicate ben 31! Mi limito ad elencare solo alcuni assiomi che, come verificato in [7], equivalgono alle prime 12 proprietà.

A_1 xCy implica yCx
A_2 non esiste una regione contenente tutte le altre,
A_3 per ogni x ed y esiste z tale che zCx e zCy,
A_4 per ogni x, xCx (proprietà riflessiva)
A_5 se $C(x) = C(y)$ allora $x = y$,
A_6 per ogni regione z esistono $x \leq z$ ed $y \leq z$ tali che x non è connesso con y.

Definizione 3. Chiamiamo *spazio di connessione* una struttura $(\mathfrak{R}, C)$ che verifica A_1-A_6.

A_1 e A_4 esprimono la proprietà simmetrica e riflessiva della relazione di connessione. A_2 mostra che l'intero spazio non viene considerato come regione e che quindi probabilmente l'idea di regione di Whitehead comprende le sole regioni limitate (ciò spiega anche l'esigenza di A_6). A_5 permette di provare la proprietà antisimmetrica di $\leq$ e quindi il fatto che $\leq$ sia una relazione d'ordine. Da A_6 segue che ogni regione contiene propriamente altre regioni e che quindi non esistono atomi-punto nelle strutture di connessione. Accanto alla nozione di connessione definiamo quella di *sovrapposizione*. Diciamo che due regioni x ed y sono *sovrapposte* se esiste una regione z contenuta sia in x che in y. Per indicare che x è sovrapposto ad y scriveremo xSy, inoltre indichiamo con $S(x)$ l'insieme delle regioni che si sovrappongono ad x. E' immediato provare che se y si sovrappone ad x allora y è connesso ad x. Tuttavia in generale non vale il viceversa e la relazione di connessione (di carattere topologico) non coincide in generale con quella di sovrapposizione (legata alla sola nozione di ordinamento).

Naturalmente risulta fondamentale la ricerca di modelli matematici plausibili per le strutture di inclusione e per quelle di connessione. Come viene fatto nel caso delle geometrie non euclidee, in cui si sono costruiti modelli a partire dalla usuale geometria euclidea, cercheremo modelli di geometria senza punti all'interno della usuale geometria euclidea. Pertanto partiamo dallo spazio euclideo a tre dimensioni ed in questo ambiente cerchiamo ragionevoli candidati per rappresentare la nozione di "regione dello spazio". La nozione di chiuso regolare sembra essere adeguata.[26]

Definizione 4. Chiamiamo *regione* ogni chiuso regolare dello spazio euclideo.

Si noti che la classe degli insiemi chiusi regolari dello spazio euclideo è una algebra di Boole senza atomi e quindi costituisce una

---

[26] Nelle pagine successive utilizzeremo tale nozione per definire quella di figura geometrica.

base elegante per la costruzione di modelli. Volendo procedere più in generale, è possibile riferirsi ad un qualunque spazio topologico.

Definizione 5. Uno spazio di inclusione ($\Re$, $\leq$) si dice *canonico* se $\Re$ è un insieme di chiusi regolari limitati e non vuoti di uno spazio topologico e $\leq$ è la relazione di inclusione. Uno spazio di connessione ($\Re$, C) è detto *canonico* se $\Re$ è un insieme di chiusi regolari limitati e non vuoti di uno spazio topologico e C è la relazione binaria in $\Re$ definita ponendo

$$XCY \Leftrightarrow c(X) \cap c(Y) \neq \emptyset.$$

3. Un approccio metrico

Una alternativa a quanto proposto da Whitehead è quella di riferirsi ad uno dei più importanti approcci alla geometria: quello metrico (si veda in proposito il bel libro di E. V. Huntington). Ricordiamo che uno *spazio metrico* è definito da un insieme non vuoto M, i cui elementi sono chiamati *punti*, e da una funzione $d : M \rightarrow [0,\infty)$, chiamata *distanza*, tale che,

(1) $d(x, x) = 0$, (3) $d(x,y) \leq d(x, z) + d(z, y)$

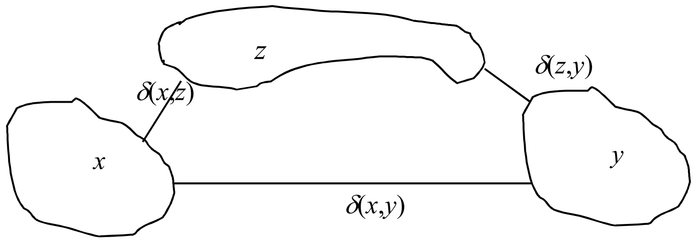
(2) $d(x,y) = d(y,x)$, (4) $d(x,y) = 0 \Rightarrow x = y$,

per ogni x, y, z in M. Se sono verificate solo le proprietà (1), (2) e (3) la struttura (M, d) sarà chiamata *spazio pseudo-metrico*. Ad ogni spazio pseudo-metrico è associato uno spazio metrico (M',d) che è il quoziente di M modulo la relazione di equivalenza $\{(x,x') : d(x,x') = 0\}$. Più precisamente, avremo che $M' = \{[x] : x \in M\}$ dove, per ogni $x \in M$, $[x] = \{x' : d(x,x') = 0\}$. Inoltre poniamo $d([x],[y]) = d(x,y)$.

Ora, se si vuole trovare una nozione analoga a quella di spazio metrico che però si riferisca alle regioni dello spazio, conviene ammettere per un momento che le regioni siano sottoinsiemi di uno spazio metrico. Ciò consente di fare riferimento alla nota definizione di distanza tra due insiemi x ed y ottenuta ponendo

$$\delta(x,y) = Inf\{d(P,Q) : P \in x,\ Q \in y\}. \qquad (3.1)$$

Naturalmente δ non è una metrica poiché da $\delta(x,y) = 0$ non segue che $x = y$. Non è nemmeno una pseudo-metrica non essendo verificata la disuguaglianza triangolare. Per rendersene conto basta osservare la seguente figura:



dove appare evidente che $\delta(x,y) > \delta(x,z)+\delta(z,y)$. D'altra parte l'intuizione suggerisce che se si considerano regioni "molto piccole" allora δ deve verificare, almeno approssimativamente, le proprietà di uno spazio metrico. Facciamo allora entrare in gioco un'altra nozione che in qualche modo sia capace di dare conto del concetto di "regione piccola": la nozione di diametro. Ricordiamo che il diametro $|x|$ di un insieme x non vuoto di punti è definito tramite

$$|x| = Sup\{d(P,Q) : P\in x,\ Q\in x\}. \qquad (3.2)$$

Possiamo ora proporre una diseguaglianza che in qualche modo gioca il ruolo della diseguaglianza triangolare. Infatti, come non è difficile verificare, le funzioni definite tramite (3.1) e (3.2) verificano la seguente relazione

$$\delta(x,y) \le \delta(x,z) + \delta(z,y) + |z|$$

la quale, se si considerano solo regioni con diametri trascurabili, viene a coincidere proprio con la diseguaglianza triangolare. Ciò suggerisce la seguente definizione (si veda [6]).

Definizione 1. Una struttura $(\mathfrak{R}, \le, \delta, |\ |)$ prende il nome di *spazio pseudo-metrico senza punti* (in breve *p-p-m*-spazio) se $\le$ è una relazione d'ordine in $\mathfrak{R}$ e le funzioni

$$\delta : \mathfrak{R}\times\mathfrak{R} \to [0,\infty) \text{ e } |\ | : \mathfrak{R} \to [0,\infty)$$

Verificano gli assiomi:

A_1 $x \le y \Rightarrow |x| \le |y|$ (crescenza del diametro),

A_2 $\delta(x,y) = \delta(y,x)$ (simmetria),

A_3 $x_1 \ge x_2 \Rightarrow \delta(x_1,y) \le \delta(x_2,y)$ (decrescenza della distanza)

A_4 $\delta(x,x) = 0$,

A_5 $\delta(x,y) \le \delta(x,z) + \delta(z,y) + |z|$ (diseguaglianza triangolare).

Gli elementi di $\mathfrak{R}$ sono chiamati *regioni*, la relazione $\le$ *relazione di inclusione*, δ *funzione distanza*, $|\ |$ *funzione diametro.*

Chiamiamo *canonico* un *p-p-m*-spazio $(\mathfrak{R}, \le, \delta, |\ |)$ tale che $\mathfrak{R}$ sia una classe dei chiusi regolari limitati (non vuoti) di uno spazio metrico, $\leq$ sia la relazione di inclusione e le funzioni δ e $|\,|$ siano definite tramite (3.1) e (3.2).

4. Modelli "concreti": una ricerca aperta

Sarebbe utile tentare di "costruire" ambienti geometrici che non solo partano dall'idea di regione ma che in un certo senso permettano di rappresentare attività concrete di misurazione. Non è chiaro come questo si possa fare ma mostrerò un paio di possibili strade.

Supponiamo di partire solo da una funzione diametro $|\,|$ e da una relazione di inclusione $\subseteq$ assunte entrambe come primitive. Allora possiamo definire la distanza tra due regioni x ed y come una misura del percorso minimo che congiunge x con y. Per essere più precisi consideriamo una struttura del tipo $(\Re, |\,|, \subseteq)$ dove

- $\Re$ è un insieme i cui elementi sono chiamati *regioni,*
- $|\,| : \Re \to [0,\infty)$ una funzione chiamata *funzione diametro*
- $\subseteq$ una relazione d'ordine in $\Re$ chiamata *inclusione*.

Chiamiamo *percorso* tra x ed y una successione di regioni $x_1,\ldots,x_n$ tali che ogni regione si sovrappone con la successiva, x si sovrappone a x_1 ed y si sovrappone a x_n. Possiamo allora porre

$$\delta(x,y) = Inf\{|x_1|+\ldots+|x_n| : x_1,\ldots,x_n \text{ percorso tra } x \text{ ed } y\}. \qquad (3)$$

Non è difficile mostrare che, fatte alcune semplici ipotesi sulla struttura $(\Re, |\,|, \subseteq)$, in questo modo si ottiene un *p-p-m*-spazio.

Una via leggermente diversa si ottiene introducendo la nozione di "movimento" al posto di quella di diametro. Si potrebbe partire da una struttura del tipo $(\Re, G, \subseteq)$ dove G è un gruppo di automorfismi di $(\Re, \subseteq)$, cioè di funzioni invertibili in $\Re$ che conservano la relazione $\subseteq$ e la cui inversa conserva ancora la relazione $\subseteq$. L'idea intuitiva è che G dovrebbe rappresentare gli spostamenti rigidi. Chiamiamo *equivalenti* due regioni x ed y tali che $g(x) = y$ per una opportuno $g\in G$. Detto questo potremmo fissare una regione u (*regolo unitario*) e poi definire δ utilizzando la formula (3) ma:

1. limitandoci a percorsi formati da regioni tutte equivalenti ad u
2. attribuendo diametro 1 a tali regioni.

In questo modo si otterrebbero solo misure intere, per avere misure migliori dovrebbe essere possibile dividere un regolo in più parti. In [15], Schmidt propone una interessante teoria non molto lontana da tale idea. Tuttavia, specialmente se si vuole tenere conto degli aspetti didattici dell'insegnamento della geometria, le ricerche in proposito sono ancora tutte da sviluppare.

5. Creare i punti tramite i processi di astrazione

Passiamo ora alla nozione che maggiormente ci interessa, quella di punto. Whitehead definisce i punti tramite i *processi di astrazione* intesi come processi che consentono una approssimazione-costruzione sempre migliore ad un ente "ideale". L'idea-guida è che un punto, un segmento od una superficie siano rispettivamente il "limite" di una successione di regioni sempre più piccole, sempre più sottili, sempre meno spesse.

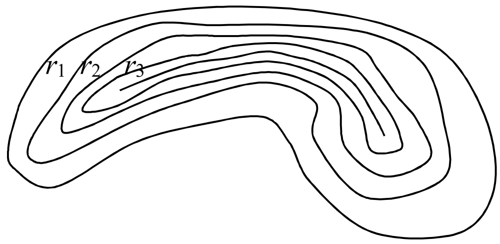
Definizione 1. Dato uno spazio di inclusione $(\Re,\leq)$, un *processo di astrazione* è una successione $(r_n)_{n\in N}$ di regioni tale che:

(i) $(r_n)_{n\in N}$ è strettamente decrescente ;

(ii) non esiste una regione contenuta in tutte le regioni di $(r_n)_{n\in N}$.

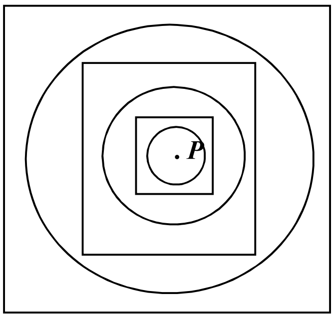
Denotiamo con PA l'insieme dei processi di astrazione.

Intuitivamente si deve immaginare che un processo di astrazione $(r_n)_{n\in N}$ conduca a definire la figura geometrica $\bigcap_{n\in N} r_n$ e che tale figura sia di dimensione inferiore a quella delle regioni. Ecco un esempio di processo di astrazione che definisce una curva:



Ovviamente si deve tenere ben presente che, non essendo le regioni insiemi di punti, la nozione insiemistica di intersezione non ha senso. D'altra parte la condizione ii) comporta che, nell'ambiente "concreto" delle regioni, non esiste nemmeno l'estremo inferiore di un processo di astrazione e che quindi tali processi definiscono qualcosa di completamente nuovo rispetto a tale ambiente.

Il secondo passo è quello di identificare processi di astrazione che conducono allo stesso ente. Ad esempio, dato un punto P, indichiamo con s_n la sfera di centro P e diametro $1/n$ e con c_n il cubo di centro P e lato $1/n$. Allora la nostra intuizione suggerisce che $(s_n)_{n\in N}$ e $(s_n)_{n\in N}$ rappresentano entrambi il punto P.



Pertanto è necessario definire in qualche modo una relazione di equivalenza tra processi di astrazione.

Definizione 2. Diciamo che il processo $(s_n)_{n\in N}$ *copre* il processo $(t_n)_{n\in N}$ e scriviamo $(s_n)_{n\in N} \geq (t_n)_{n\in N}$ se per ogni regione s_i esiste una regione t_j contenuta in s_i.

Intuitivamente $(s_n)_{n\in N}$ copre $(t_n)_{n\in N}$ se e solo se $\bigcap_{n\in N} s_n$ contiene $\bigcap_{n\in N} t_n$. Infatti mettiamoci nell'ipotesi che tali regioni siano insiemi di punti e supponiamo che $(s_n)_{n\in N}$ copra $(t_n)_{n\in N}$. Allora detto P un punto di $\bigcap_{n\in N} t_n$ per ogni regione s_i esiste t_j è tale che $t_j \subseteq s_i$. Poiché $P\in t_j$ risulta quindi $P\in s_i$. Ne segue che P appartiene a tutte le regioni s_i e quindi appartiene all'intersezione $\bigcap_{n\in N} s_n$.

La relazione di copertura è un preordine in PA e quindi effettuando un opportuno quoziente di $(PA, \leq)$ otteniamo un insieme ordinato. Più precisamente si procede al modo seguente.

Proposizione 3. Definiamo in PA la relazione $\equiv$ ponendo

$$(s_n)_{n\in N} \equiv (t_n)_{n\in N} \Leftrightarrow (s_n)_{n\in N} \leq (t_n)_{n\in N} \text{ e } (s_n)_{n\in N} \geq (t_n)_{n\in N}.$$

Allora $\equiv$ è una relazione di equivalenza. Poniamo $EG = PA/\equiv$ e definiamo in EG la relazione $\leq$ ponendo

$$[(s_n)_{n\in N}] \leq [(t_n)_{n\in N}] \Leftrightarrow (s_n)_{n\in N} \leq (t_n)_{n\in N}.$$

Allora $(EG,\leq)$ è un insieme ordinato.

Intuitivamente i due processi di astrazioni $(s_n)_{n\in N}$ e $(t_n)_{n\in N}$ sono equivalenti se $\bigcap_{n\in N} s_n = \bigcap_{n\in N} t_n$. La seguente definizione è implicita nella proposta di Whitehead.

Definizione 4. Chiamiamo *elemento geometrico* (astratto) ogni elemento di *EG* cioè ogni classe completa di equivalenza nell'insieme dei processi di astrazione. Chiamiamo *inclusione* la relazione d'ordine in *EG*.

Se ci si riferisce allo spazio euclideo tridimensionale, gli elementi geometrici astratti sono le superfici, le linee ed i punti. Possiamo definire i punti come quei particolari elementi geometrici che, in accordo con la definizione di Euclide, *"non hanno parti"*.

Definizione 5. Chiamiamo *punto* un elemento geometrico che non include propriamente nessun altro elemento geometrico.

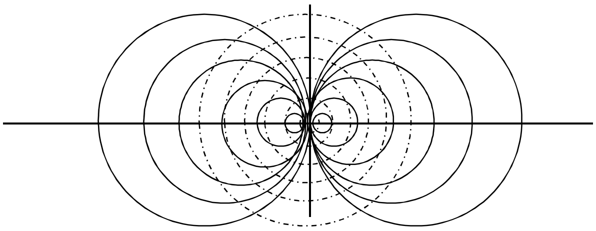
E' importante osservare che, in termini logici, contrariamente a quanto avviene di solito, un punto è un oggetto del secondo ordine mentre una regione è un oggetto del primo ordine.

Si noti che un punto non deve essere considerato come il risultato di un processo continuo di approssimazione ma, in un certo senso, come il processo di approssimazione. In tale senso, l'infinito attuale non viene coinvolto in quanto il processo non deve essere necessariamente immaginato come terminato. Allo stesso modo, se si definiscono i numeri reali tramite le successioni di Cauchy, allora un numero reale è una successione di Cauchy (il processo), non il limite di tale successione (il risultato del processo). Tale approccio, abbastanza generale in matematica, si basa sull'equazione

(*spazio di nuovi enti*) = *processi* + (*equivalenza tra processi*).

6. Ma in questo modo si creano troppi punti

Nel caso delle strutture di inclusione la definizione di punto di Whitehead non è soddisfacente. Ad esempio, riferendoci al piano euclideo, possiamo considerare tre classi astrattive. La prima è definita dalla successione $(a_n)_{n\in N}$ di cerchi di centro $(-1/n,0)$ e raggio $1/n$, la seconda dalla successione di cerchi $(b_n)_{n\in N}$ di centro $(1/n,0)$ e raggio $1/n$ ed infine la terza $(c_n)_{n\in N}$ dalla successione di cerchi di centro l'origine e raggio uguale a $1/n$.



Allora, poiché $(a_n)_{n\in N}$ e $(b_n)_{n\in N}$ non si coprono, gli enti geometrici corrispondenti $0^- = [(a_n)_{n\in N}]$ e $0^+ = [(b_n)_{n\in N}]$ sono diversi tra loro. Ancora diverso è l'ente geometrico $0 = [(c_n)_{n\in N}]$. In definitiva vengono ad essere definiti tre distinti elementi geometrici con 0 "contenente" propriamente 0^- e 0^+. Questo significa che 0 non può essere visto come punto, cosa questa abbastanza lontana dall'intuizione. Naturalmente una tale "patologia" potrebbe essere interessante e mediante essa si potrebbe rendere rigoroso l'uso dei simboli 0^-, 0^+, uso che spesso viene fatto in analisi matematica. Come vedremo nel seguito, è naturale sviluppare l'idea di uno spazio in cui i punti classici si "spezzano" in più punti.

Si noti che un altro tipo di variazione che permette di ottenere nuovi punti si ottiene ammettendo come regioni anche regioni non limitate. In tale modo si potrebbero ottenere come punti anche "punti all'infinito".Qualche cosa di simile viene fatto nell'analisi non standard[27] dove esistono numeri infiniti ed infinitesimi ed ogni numero reale esplode in una nuvola di numeri non standard infinitamente vicini tra loro.

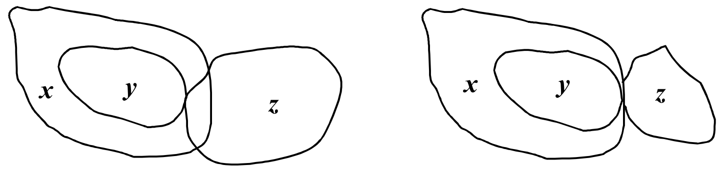
Tali difficoltà non sono invece presenti nel caso delle strutture di connessione. Probabilmente è questo che ha portato Whitehead a passare dalle strutture di inclusione a quelle di connessione. Lo strumento utilizzato da Whitehead è la nozione di inclusione tangenziale.

Definizione 1. Data una struttura di connessione, chiamiamo *inclusione non tangenziale*, la relazione « definita ponendo

$$x \ll y \iff C(x) \subseteq S(y).$$

[27] Si veda nel Capitolo 3 il paragrafo sui razionali non-standard.

In altri termini $x \ll y$ vale se una regione z non può connettersi ad x senza sovrapporsi ad y. Invece x è inclusa tangenzialmente in y se x è inclusa in y ed esiste una regione z che si connette ad x senza sovrapporsi ad y. Ecco un esempio di inclusione non tangenziale ed uno di inclusione tangenziale.



Tale nozione permette di definire in modo più adeguato i processi di astrazione.

Definizione 2. Un *processo di astrazione* in uno spazio di connessione $(\Re,C)$ è una successione di regioni $(r_n)_{n\in N}$ tale che:

a) $(r_n)_{n\in N}$ sia decrescente rispetto a $\ll$

b) non esiste nessuna regione che sia contenuta in ogni r_n.

La nozione di elemento geometrico e di punto si definiscono poi come nel paragrafo precedente. Un tale modo di procedere elimina le difficoltà esposte all'inizio del paragrafo. Infatti le classi $(a_n)_{n\in N}$ e $(c_n)_{n\in N}$ non sono processi di astrazione in quanto le regioni che le compongono sono incluse tangenzialmente.

Si osservi che nel modello canonico definito dai regolari (non necessariamente limitati) le nozioni di punto e di figura geometrica non coincidono con quelle usuali della geometria euclidea. Infatti in tale caso ci si ritrova anche con enti geometrici "all'infinito". Per esempio, consideriamo la successione $(r_n)_{n\in N}$ dei semipiani $r_n = \{(x,y) \in R^2 : x \geq n\}$. Allora tale successione è un processo di astrazione che rappresenta una "figura geometrica" che non ha punti al finito.

7. Dalla topologia alla geometria: ovali, segmenti e distanze

La relazione di connessione è di carattere topologico e può pertanto esprimere solo nozioni di natura topologica. Per poter passare ad un livello più propriamente geometrico è possibile procedere in diversi modi. Whitehead in *Processo e Realtà* per poter definire la nozione di segmento postula l'esistenza di una particolare classe di regioni, chiamate *ovali*. In questo articolo rinunciamo a descrivere gli assiomi relativi agli ovali. L'idea di ovale sembra comunque corrispondere a quella di insieme convesso dello spazio euclideo, cioè insieme X tale che se due punti P e Q appartengono ad X allora tutti i punti del segmento PQ appartengono ancora ad X. Da notare che la classe dei convessi è un sistema di chiusura, cioè l'intersezione di una famiglia di convessi è ancora un convesso. In particolare ha senso quindi parlare di insieme convesso *generato* da due punti P e Q come intersezione di tutti i convessi contenenti P e Q. Naturalmente tale insieme coincide con il segmento PQ congiungente P con Q. Ciò suggerisce la seguente definizione.

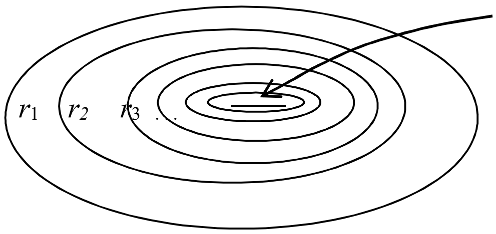
Definizione 1. Dati due punti P e Q, il *segmento tra* P e Q è l'elemento geometrico

a) contenente P e Q,

b) rappresentabile da un processo di astrazione costituito da ovali

c) minimale rispetto alle due condizioni di cui sopra.

Nella figura seguente si ottiene un segmento di estremi P e Q



Ovviamente è possibile definire in modo analogo anche la nozione di figura triangolare e di tetraedo. Una volta definita la nozione di segmento non ci sono difficoltà a definire tutte le altre nozioni della geometria affine.

Naturalmente volendo definire la nozione di modello canonico di spazio affine senza punti, non possiamo più considerare gli spazi topologici ma dobbiamo riferirci a spazi in cui è definibile la nozione di convessità.

E' un problema aperto la ricerca di un adeguato sistema di assiomi da imporre alla nozione di ovale in modo che la risultante nozione di segmento permetta di definire la geometria euclidea.

Più facile è la definizione di concetti geometrici nel caso dei *p-p-m*-spazi. Infatti come è noto l'intero edificio della geometria euclidea può essere definito in termini di nozioni metriche (si veda ad esempio [2]). Basta allora mostrare come sia possibile associare ad ogni *p-p-m*-spazio uno opportuno spazio metrico. Naturalmente per fare questo dobbiamo ancora una volta dare una buona definizione di punto. Ora siamo spesso abituati a dire che un punto (un punto materiale) si può rappresentare come una regione (un corpo) "molto piccola" intendendo dire con tale espressione che quanto più piccola è la regione (il corpo) cui ci si riferisce tanto più precisa risulta la rappresentazione. Ciò suggerisce la seguente definizione.

Definizione 2. Chiamiamo *processo di astrazione* in un *p-p-m-spazio* $(\Re, \leq, \delta, |\,|)$ una successione decrescente di regioni $(r_n)_{n\in N}$ tale che $lim_{n\to\infty} |r_n| = 0$. Indichiamo con *PA* l'insieme dei processi di astrazione. La "distanza" tra due processi di astrazione è data da:

$$d'((x_n)_{n\in N}, (y_n)_{n\in N}) = lim_{n\to\infty}\, \delta(x_n, y_n).$$

(PA,d') è uno spazio pseudometrico ma non è, in generale, uno spazio metrico. Come è noto, possiamo ottenere uno spazio metrico considerando un opportuno quoziente di (M', d).

Proposizione 3. Definiamo in *PA* la relazione $\equiv$ ponendo

$$(x_n)_{n\in N} \equiv (y_n)_{n\in N} \Leftrightarrow d'((x_n)_{n\in N}, (y_n)_{n\in N}) = 0.$$

Allora $\equiv$ è una relazione di equivalenza. Definiamo in $M = PA/\!\equiv$ la funzione d ponendo

$$d([(x_n)_{n\in N}],[(y_n)_{n\in N}]) = d'((x_n)_{n\in N}, (y_n)_{n\in N}) = lim_{n\to\infty}\, \delta(x_n,y_n).$$

Allora (M, d) è uno spazio metrico.

Definizione 4. Diciamo che (M,d) è la *spazio metrico associato a* $(\Re, \leq, \delta, |\,|)$ e chiamiamo *punto* ogni elemento di M.

Ne segue che un punto è una classe completa di equivalenza

$$P = [(x_n)_{n\in N}] = \{<y_n> \mid d'(<x_n>,<y_n>) = 0\}.$$

E' un problema aperto trovare un sistema di assiomi per $(\Re, \leq, \delta, |\,|)$ in modo che lo spazio metrico associato (M,d) sia isomorfo allo spazio metrico euclideo a tre dimensioni.

8. Tornando alla questione del gatto

Torniamo alla questione da cui siamo partiti, cioè se abbia senso parlare di un punto dall'aspetto di gatto. Essendoci liberati dal pregiudizio "punto = elemento primitivo", possiamo tentare ora di dare una risposta positiva alla questione: basta modificare opportunamente le definizioni di punto che abbiamo esposto. Ad esempio, consideriamo il *p-p-m*-spazio canonico individuato dalla classe di tutti i chiusi regolari dello spazio metrico a tre dimensioni. Allora potremmo "rafforzare" la relazione di equivalenza già proposta per i processi di astrazione e richiedere, per esempio, che due processi di astrazione per essere equivalenti oltre che coprirsi a vicenda debbano essere costituiti di regioni "con la stessa forma". In tale caso, ad esempio, le successioni $(s_n)_{n\in N}$ e $(c_n)_{n\in N}$ di sfere e di cubi di centro in un punto P di cui abbiamo parlato nel primo paragrafo non sarebbero da considerarsi equivalenti e verrebbero a definire due diversi punti; uno sferico ed uno cubico. D'altra parte è evidente che ogni rafforzamento della relazione di equivalenza corrisponde ad un aumento del numero delle classi di equivalenza e quindi ad un aumento del numero dei punti del nostro spazio.

Non è difficile formalizzare un tale modo di procedere e definire la nozione di *"punto con una data forma"*.

Definizione 1. Chiamiamo *similitudine di rapporto k* in un *p-p-m*-spazio una trasformazione f che conserva l'inclusione e tale che, per ogni coppia r ed s di regioni

$$d(r,s) = k \cdot d(f(r),f(s)) , \quad |r| = k \cdot |f(r)|.$$

Chiamiamo *simili* due regioni che sono una l'immagine dell'altra tramite una similitudine.

Poiché l'insieme delle similitudini costituisce un gruppo, la relazione di similitudine tra regioni è di equivalenza. Chiamiamo *forma* una classe completa di equivalenza rispetto la relazione di similitudine.

Definizione 2. Chiamiamo *processo di astrazione uniforme* un processo di astrazione in cui tutte le regioni hanno la stessa forma. Indichiamo con *RPF* l'insieme dei processi di astrazione uniformi e chiamiamo *equivalenti* due elementi $(r_n)_{n\in N}$ e $(t_n)_{n\in N}$ in *RPF* se:

- si coprono a vicenda

- le regioni di $(r_n)_{n\in N}$ hanno la stessa forma delle regioni di $(t_n)_{n\in N}$.
Chiamiamo *punto con forma* una classe completa di equivalenza determinata in *RPF* da tale relazione di equivalenza.

Con tale scelta ogni punto (inteso come sola posizione nello spazio), esplode in una infinità di punti dalle forme più diverse. Naturalmente tra i punti con forma esisteranno anche punti dall'aspetto di gatto.

Per quanto riguarda la questione dei punti che ruotano intorno a se stessi, l'intuizione ci dice che tale idea per astrazione dal considerare un qualunque corpo che ruoti intorno a se stesso e che sia abbastanza piccolo da poter essere assimilato ad un punto. Appare quindi naturale definire un punto che ruoti intorno a se stesso come limite di una sequenza di regioni sempre più piccole che ruotano intorno a se stesse.

Concludendo, non sembra poi tanto assurdo pensare ad un punto dall'aspetto di gatto ed affermare che tale punto ruota verso noi sorridendo.

> *- Hai ragione, - disse il Gatto - e questa volta svanì adagio adagio cominciando con la fine della coda e finendo col sorriso, il quale rimase per qualche tempo sul ramo, dopo che tutto era svanito.*
> *- Curioso! ho visto spesso un gatto senza sorriso; - osservò Alice, - mai un sorriso senza Gatto. È la cosa più strana che mi sia capitata!* (L. Carroll, *Alice nel paese delle meraviglie*).

Bibliografia

[1] L. Biacino, G. Gerla, Connection structures, *Notre Dame Journal of Formal Logic*, 32 (1991) 242-247.
[2] Blumental L. M., *Theory and applications of Distance Geometry*. Chelsea Publishing Company, N. Y. 1970.
[3] S. Borgo, N. Guarino, C. Masolo, *Stratified Ontologies: The Case of Physical Objects*, ECAI-96 Workshop on Ontological Engineering.
[4] S. Borgo, N. Guarino, C. Masolo, A Naive Theory of Space and Matter, in P. Amsili, M. Borillo, and L. Vieu (eds.), *Time, Space and Movement: Meaning and Knowledge in the Sensible World* (Proceedings of the 5th International Workshop), Toulouse: COREP, 1995, pp. 29–32.
[5] R. Casati, A. Varzi, *Buchi e altre superficialità*, Garzanti

(1996).
[6] G. Gerla, Pointless metric spaces, *The Journal of Symbolic Logic*, 55 (1990), 207-219.
[7] G. Gerla, R. Tortora, La relazione di connessione in A. N. Whitehead: aspetti matematici, *Epistemologia*, 15 (1992) 341-354.
[8] G. Gerla, Distances, diameters and verisimilitude of theories, *Archive for Mathematical Logic*, 31 (1992) 407-414.
[9] G. Gerla, Pointless geometries, in *Handbook of Incidence Geometry*, F. Buekenhout and W. Kantor (eds) 1994 North-Holland.
[10] G. Gerla., R. Volpe, Geometry without points, *The American Math. Monthly*, 92 (1985) 707-711.
[11] G. Gerla, R. Tortora, Dissezioni e intersezioni di regioni in A. N. Whitehead, *Epistemologia*, 19 (1996) 289-308.
[12] E.V.Huntington, Postulates for abstract geometry, *Mathematische Annalen*, 73 (1913) 522-559.
[13] N. I. Lobačevskij, *Nuovi principi della geometria*, Ed. Boringhieri (1955).
[14] W. Noll, The geometry of contact, separation, and reformation of continuous bodies, *Arch. Ration. Mech. Anal.*, 122 (1993) 197-212.
[15] H. J. Schmidt, *Axiomatic Characterization of Physical Geometry*, Lecture Notes in Physics, Springer-Verlag, Berlin Heidelberg 1979.
[16] A. N. Whitehead, *An Inquiry Concerning the Principles of Natural Knowledge*, Cambr. Univ. Press, Cambrige 1919.
[17] A. N. Whitehead, *The concept of Nature*, Cambr. Univ. Press, Cambrige 1920.
[18] A. N. Whitehead, *Process and Reality*, The Macmillan Co., New York 1929.

CAPITOLO 2

CRISI DELLA GEOMETRIA EUCLIDEA

Per tutto il diciassettesimo ed il diciottesimo secolo la geometria rimase, nella guerra contro l'empirismo, una fortezza inespugnabile degli idealisti. Coloro i quali credevano (come in generale si credeva sul continente) che fosse possibile una conoscenza del mondo reale certa ed indipendente dall'esperienza non avevano che da indicare la geometria: soltanto un pazzo avrebbe messo in dubbio la sua validità, e soltanto uno sciocco ne avrebbe negato il riferimento oggettivo. (Bertrand Russell)

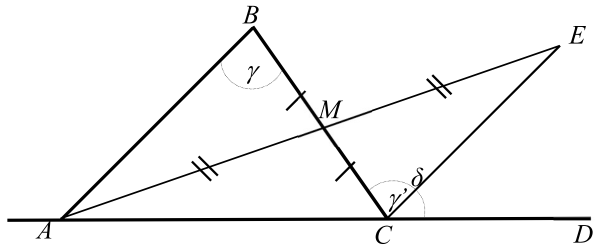
1. Crisi del carattere assoluto della geometria

La geometria euclidea restò al centro della matematica fino agli inizi del 1800 quando, ad opera di matematici come Lobačevskij e Bolyai si svilupparono le geometrie non euclidee, cioè geometrie che negavano validità al quinto postulato. Prima di parlare delle geometrie non euclidee, esaminiamo più da vicino il significato del quinto postulato. Per prima cosa mostriamo che dagli assiomi di Euclide escluso il quinto è possibile comunque derivare che data una retta r ed un punto P fuori da essa esiste almeno una retta parallela ad r passante per P. Il quinto postulato serve invece per dimostrare che tale retta è unica. Pur non volendo esporre le dimostrazioni complete, vediamo i passi fondamentali per provare l'esistenza della parallela e cominciamo con il Teorema dell'angolo esterno.

Proposizione 1.1. Un angolo esterno di un triangolo è sempre strettamente maggiore dei due angoli interni non adiacenti.

Dim. Consideriamo un triangolo ABC e proviamo ad esempio che l'angolo esterno $\delta = BCD$ è maggiore dell'angolo $\gamma = ABC$. A tale scopo denotiamo con M il punto medio del segmento BC e prolunghiamo il segmento AM in un segmento AE in modo che AM sia uguale a ME. Allora i due triangoli ABM e MEC sono uguali avendo i due angoli in M uguali in quanto opposti al vertice e due lati uguali per costruzione. In particolare γ sarà uguale all'angolo

$\gamma' = BCE$. D'altra parte tale angolo, essendo una parte di δ, risulta minore di δ. In conclusione γ è minore di δ.

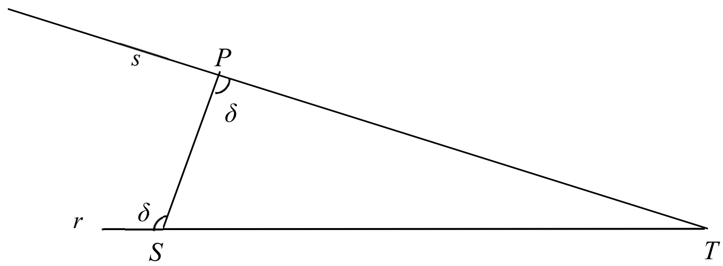


□

Naturalmente la dimostrazione che abbiamo dato diventa rigorosa solo se prima si è provato che tutte le costruzioni fatte sono rese possibili dai postulati di Euclide (escluso il quinto). Ad esempio deve essere prima provato che esiste il punto medio di un segmento, che due angoli opposti al vertice sono uguali e che valgono i criteri di uguaglianza dei triangoli. Discorso analogo vale per tutte le dimostrazioni di carattere euclideo presenti in questo capitolo.

Proposizione 1.2. Data una retta r ed un punto P non appartenente ad r esiste almeno una retta per P parallela ad r.

Dim. Sia S un punto qualsiasi della retta r, tracciamo la retta per



T

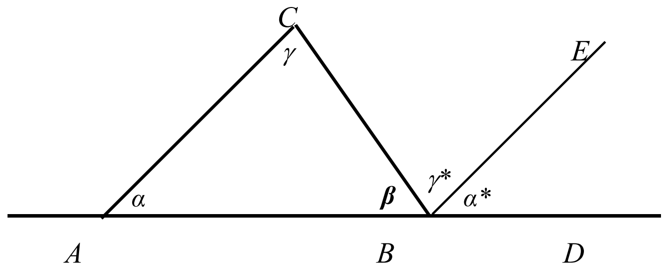
S e per P ed indichiamo con δ uno degli angoli tra r ed il segmento SP. Tracciamo poi una retta s per P che formi con SP l'angolo δ. Allora se per assurdo le rette r ed s si incontrassero in un punto T, nel triangolo SPT avremmo l'angolo esterno in S uguale

all'angolo interno non adiacente in *P*, in contrasto con quanto dimostrato nella proposizione precedente. Quindi *s* non incontra *r*
. □

Se si accetta il quinto postulato allora è possibile dimostrare un teorema più generale del teorema dell'angolo esterno e questo anche in maniera più semplice.

Teorema 1.3. Un angolo esterno è uguale alla somma degli angoli interni non adiacenti. Ne segue che la somma degli angoli interni di un triangolo è un angolo piatto.

Dim. Dato il triangolo di vertici *A*, *B*, *C* tracciamo da *B* la



parallela *BE* alla retta *AC*. In tale modo si ottiene una coppia di rette parallele tagliate dalla trasversale *CB*. Per un teorema (che supponiamo di avere già dimostrato) sugli angoli alterni interni abbiamo che $\gamma = \gamma^*$ mentre per il teorema sugli angoli corrispondenti abbiamo che $\alpha = \alpha^*$. Ciò prova che l'angolo esterno *CBD* è uguale ad $\alpha + \gamma$. Ne segue anche che $\beta + \alpha + \gamma = \beta + \gamma^* + \alpha^*$ è un angolo piatto. □

Da tale teorema ovviamente segue anche il teorema dell'angolo esterno. Tuttavia Euclide preferisce effettuare la dimostrazione del teorema dell'angolo esterno che abbiamo esposto in Proposizione 1.1 probabilmente perché non utilizza il postulato delle parallele. D'altra parte l'uso di tale postulato viene rimandato negli Elementi quanto più possibile quasi a dimostrazione di qualche perplessità. Naturalmente i matematici greci erano convinti della verità del quinto postulato. Il problema è che pare fossero anche convinti che esso fosse in realtà dimostrabile utilizzando gli altri assiomi della geometria. Tale convincimento rimase comune anche a tutti i matematici successivi fino al 1800. Infatti numerosi furono i tentativi di trovare una sua dimostrazione.

Ora, nella prima metà del 1800 accade un fatto nuovo: tre differenti matematici svilupparono un nuovo tipo di geometria. Si trattava del tedesco Gauss, dell'ungherese Bolyai e del russo Lobačevskij che fecero la loro scoperta uno indipendentemente dall'altro. Punto di partenza di tali autori è una nuova visione della geometria che da essi viene considerata un ramo della fisica piuttosto che un prodotto a priori del nostro spirito. Ciò comportava che la validità degli assiomi della geometria dovesse essere convalidata o confutata dall'esperienza. Ad esempio ecco cosa scrive Gauss.

> *Secondo la mia più profonda convinzione, la teoria dello spazio ha nei confronti del nostro sapere una posizione completamente diversa da quella della pura teoria delle grandezze (aritmetica); infatti, viene assolutamente a mancare alla nostra conoscenza della prima quella completa convinzione della sua necessità (e quindi anche della sua assoluta verità), che invece inerisce alla seconda; dobbiamo umilmente ammettere che, mentre il numero è puramente un prodotto del nostro spirito, lo spazio possiede una realtà anche al di fuori del nostro spirito, alla quale noi non possiamo prescrivere le sue leggi completamente a priori.* (Lettera di Gauss a Bessel del 9 aprile 1830)

Naturalmente il considerare la geometria alla stessa stregua della fisica comportava una interpretazione delle nozioni geometriche in termini di oggetti esistenti nel mondo reale. Ad esempio il segmento congiungente due punti P e Q poteva essere assimilato al percorso di minima distanza tra P e Q, oppure alla linea percorsa da un raggio di luce che partito da P raggiunge Q. Ciò permetteva ad esempio di immaginare esperimenti capaci di stabilire se un dato triangolo avesse o meno come somma di angoli interni un angolo piatto. A questo nuovo punto di vista nei confronti della geometria corrispondeva poi anche un atteggiamento diverso nei confronti del quinto postulato. Infatti se per i matematici greci il volere dimostrare tale postulato era solo una questione di eleganza e semplicità e nessuno ne metteva in dubbio la validità, i matematici dell'ottocento accettavano anche la possibilità che esso non fosse verificato nello spazio reale. Ecco quello che dice Lobačevskij nell'introduzione al suo libro "Nuovi principi della geometria" del 1835.

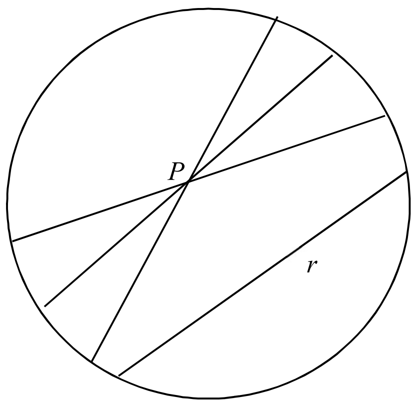
A tutti è noto che, fino ad oggi, nella geometria la teoria delle parallele è rimasta incompiuta. Gli sforzi inutili compiuti dai tempi di Euclide, per il corso di duemila anni, mi spinsero a sospettare che nei suoi stessi concetti non si racchiude ancora quella verità che si voleva dimostrare, e che può essere controllata, in modo simile alle altre leggi fisiche, soltanto dall'esperienze quali, ad esempio, le osservazioni astronomiche.

Ma nel caso che il postulato delle parallele fosse falso nello spazio reale, allora esso non poteva essere certo provato a partire dai rimanenti postulati (certamente veri). Infatti da cose vere è possibile dedurre solo cose vere.

2. Modelli di geometrie non euclidee

Lo strumento per provare in maniera inconfutabile che il quinto postulato non è derivabile dai rimanenti è quello di esibire "modelli" matematici che non verificano il quinto postulato ma che verificano tutti i postulati rimanenti. Tali modelli furono chiamati *modelli non euclidei della geometria*. Esponiamo, sommariamente e solo per darne una idea, due modelli di geometria non euclidea, quello di Klein e quello di Poincaré.

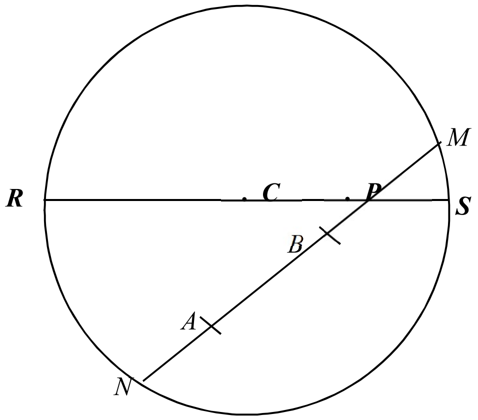
Il modello di Klein. In tale modello
- i punti sono i punti interni ad un dato cerchio S di centro C,
- le rette sono le corde del cerchio (esclusi gli estremi).



Come è illustrato nella figura, per un punto P fuori di una retta r passano infinite rette parallele ad r. Pertanto non vale l'assioma delle parallele. È immediato provare che tutti gli altri assiomi di Euclide sono verificati. Ad esempio, è evidente che per due punti passa una ed una sola retta. Inoltre la nozione di distanza di
due punti A e B (che serve per dare la nozione di eguaglianza tra segmenti) si definisce ponendo

$$d(A,B) = -log\left(\frac{MB \cdot NA}{MA \cdot NB}\right)$$

dove M ed N sono i punti di intersezione di AB con la circonferenza.



Per capire il significato di tale distanza, supponiamo ad esempio che la circonferenza abbia equazione $x^2+y^2=1$, e calcoliamoci la distanza di un punto P di coordinate $(x,0)$ dal centro C.

$$d(C,P) = -log\left(\frac{PS \cdot 1}{1 \cdot RP}\right) = -log((1-x)/(1+x)).$$

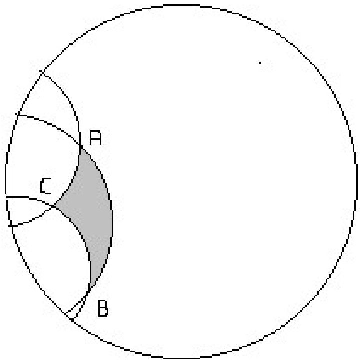
Questa formula è interessante perché mostra che:

> la distanza $d(C,P)$ tende a diventare infinita quando il punto P si avvicina al bordo del cerchio.

Ciò significa che un segmento unitario che venga spostato verso il bordo deve subire una contrazione che è tanto più forte quanto più ci si allontani dal centro. In modo equivalente possiamo dire che se, partendo dal centro del cerchio, cominciamo ad allontanarci a passi regolari, tali passi divengono sempre più piccoli senza che noi ce ne accorgiamo. Pertanto per un essere vivente all'interno del cerchio non risulta possibile uscire dal cerchio che gli apparirà, a tutti gli effetti, un universo non limitato. Ciò in contrasto con quanto appare ad un osservatore "esterno" per il quale il modello di Klein sembra occupare una parte limitata dello spazio. Possiamo immaginare una tale situazione supponendo che i punti perimetrali del cerchio esercitino una forza di repulsione verso i punti interni e che tale forza risulti tanto più forte quanto più ci si avvicina a tali bordi.

Si osservi che la nozione di uguaglianza di angoli viene definita in modo analogo a quanto fatto per l'uguaglianza di segmenti e che tale nozione non coincide con quella usuale del piano euclideo. Inoltre, nonostante le apparenze, un triangolo di tale geometria ha somma degli angoli interni minore di un angolo piatto.

Il modello di Poincaré. Consideriamo ancora come insieme di punti l'insieme dei punti interni ad una circonferenza *S* ma cambiamo la nozione di retta. Infatti chiamiamo "rette" tutte i diametri e tutte le circonferenze perpendicolari ad *S*. La nozione di lunghezza di un segmento si definisce in maniera non troppo diversa da quella del modello di Klein mentre la nozione di uguaglianza di angoli è quella usuale del piano euclideo. Ciò è interessante in quanto, come si vede nel triangolo *ABC* nella figura, permette di comprendere perché la somma degli angoli interni di un triangolo è minore di un angolo retto.



Sia il modello di Klein che quello di Poincaré permettono di dimostrare il seguente teorema:

Teorema 2.1. Esiste un modello che verifica tutti gli assiomi di Euclide tranne il quinto postulato. Pertanto il quinto postulato non può essere dimostrato a partire dai rimanenti assiomi.

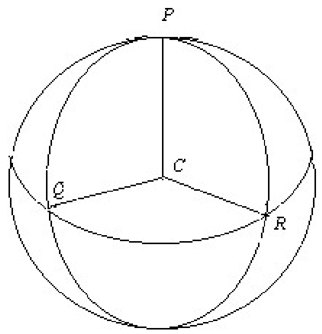
Dim. Indichiamo con *T* la teoria costituita da tutti gli assiomi proposti da Euclide escluso il quinto postulato. Allora se tale postulato fosse dimostrabile a partire da *T* esso sarebbe vero in tutti i modelli di *T*. Ma questo è impossibile poiché abbiamo trovato dei modelli di *T* che non verificano tale postulato.[1] □

---

[1] Per capire lo schema di tale ragionamento consideriamo il caso semplice in cui *T* è la teoria dei gruppi e sia α la proprietà commutativa.

3. Altre geometrie

Successivamente il matematico Riemann ed il fisico Helmholtz, uno indipendentemente dall'altro, svilupparono geometrie in cui la somma degli angoli interni di un triangolo è strettamente



maggiore di un angolo piatto. In esse non veniva negato solo il quinto postulato ma anche che un segmento possa essere prolungato a piacere. Per avere una idea di tali tipi di geometrie si osservi che nel piano euclideo i segmenti possono essere definiti come le linee di minima lunghezza congiungenti due punti assegnati. Ciò suggerisce di identificare le linee rette come lo strumento per potersi muovere con minore fatica possibile su di una superficie, e questo qualsiasi sia la superficie. Al variare delle superfici si ottengono geometrie diverse. Supponiamo ad esempio che ci si muova lungo la superficie di una sfera, allora appare ancora naturale chiamare "segmento" la linea più breve sopra tale superficie che congiunga due punti dati. È questo il punto di vista di un capitano di una nave che si muova sulla superficie terrestre e che, ovviamente, deve scegliere la rotta più breve per raggiungere la sua meta. Ora, se tale sfera ha centro C, si dimostra che tra le linee congiungenti due punti P e Q la più corta è quella che si ottiene intersecando la superficie della sfera con il cerchio di centro C e passante per P e Q. Naturalmente dei due archi per P e Q si deve scegliere il più corto è ciò crea un piccolo problema quando i due punti P e Q sono diametralmente opposti. Allora è

Allora α non può essere un teorema di T perché in tale caso tutti i gruppi sarebbero commutativi. In altre parole il mostrare un esempio di gruppo che non è commutativo mostra l'indipendenza di α da T.

opportuno considerare solo una parte della superficie sferica in modo che non vi siano punti diametralmente opposti.

Trascurando i particolari tecnici, è immediato rendersi conto che il quinto postulato non vale per tale geometria. Si consideri ad esempio la figura dove i tre punti P, Q ed R sono tali che i piani PCQ e PCR si intersecano in una retta PC ortogonale al piano CQR. Allora nel triangolo PQR gli angoli in Q ed R sono retti e ciò è in contrasto con il quinto postulato di Euclide. E' anche interessante osservare che il Teorema di Pitagora non vale in tale modello. Basta sempre riferirsi al triangolo con tre angoli retti. In tale triangolo i lati sono uguali, supposto che misurino $a \neq 0$, se valesse il Teorema di Pitagora avremmo che $a^2 = 2a^2$ da cui, dividendo per a^2 avremmo che $1 = 2$.

E facile vedere che in tutti i triangoli la somma degli angoli di tale triangolo è comunque maggiore di un angolo piatto.

Come abbiamo già detto, se si considerano superfici differenti dalla superficie sferica e si definiscono i segmenti come le linee di minima distanza (dette geodetiche), allora si ottengono altri tipi di geometrie. Pertanto esistono tanti tipi di geometrie a due dimensioni quanti sono i tipi di superficie dello spazio. Inoltre, da un punto di vista matematico, non è difficile slittare di una dimensione e considerare una "superficie tridimensionale" immersa in uno spazio a quattro dimensioni. Basta considerare una equazione a quattro incognite, chiamare "spazio quadrimensionale" l'insieme delle sue radici (cioè dei punti di una superficie nello spazio quadrimensionale) e definire al solito i segmenti come le curve continue di tale superficie che siano geodetiche. A tale superficie corrisponderà una geometria a tre dimensioni che in generale risulta essere diversa da quella euclidea.

4. Crisi dell'approccio sintetico: tutto il mondo è numero

Abbiamo visto come la scoperta delle geometrie non euclidee abbia messo in crisi il convincimento del carattere assolto della geometria Euclidea. In realtà, prima ancora di tali scoperte, un primo fondamentale elemento di crisi del metodo di Euclide (se non proprio del sistema geometrico di Euclide) si manifestò con il sistematico processo di algebrizzazione della geometria. Tale processo, iniziato nella prima metà del 1300 con Nicola d'Oresme, trasformerà la geometria "sintetica" di Euclide, in cui si dimostrano teoremi e si tracciano figure, in quella che attualmente si chiama geometria "analitica" in cui tutti i problemi si riducono alla ricerca di radici di sistemi di equazioni algebriche. In un certo senso si passa dalle "dimostrazioni con figure" tipiche della geometria euclidea ai “calcoli” tipici della geometria analitica. Infatti, come è noto, la geometria analitica si ottiene quando, fissati due assi e su di essi due unità di misura, si siano identificati

- i punti del piano con le relative coordinate,
- le rette con le equazioni di primo grado,
- le coniche con le equazioni di secondo grado e, più in generale, le curve con opportune equazioni implicite o esplicite.

Allora ad ogni operazione geometrica corrisponde una operazione di carattere analitico (cioè relativa ai numeri reali). Ad esempio l'intersezione di due curve si traduce nella risoluzione di un sistema di due equazioni.

Concorsero a tale processo di algebrizzazione scienziati e filosofi come Fermat e Cartesio. In particolare è interessante esaminare il libro di Cartesio *La Geometria* che è una delle appendici del famoso *Discorso sul Metodo* del 1637. La Geometria è costituita da tre parti, di cui la prima porta il titolo *"Dei problemi che si possono costruire col solo uso di cerchi e di linee rette"*. Si deve tenere conto che il termine "costruzione" di un problema si deve intendere come "costruzione geometrica di un segmento che sia soluzione del problema" e quindi corrisponde a "risoluzione" di un problema. In questa prima parte si illustra come sia possibile elaborare un "calcolo geometrico" dei segmenti che è l'analogo geometrico della moderna teoria dei numeri reali.

> *Come l'aritmetica è composta solo di quattro, cinque operazioni, la Somma, la Sottrazione, la Moltiplicazione, La Divisione e la Estrazione delle radici, che si può considerare una specie di Divisione, così anche in Geometria, per quanto riguarda le linee che si cercano . . .*

Il brano prosegue spiegando come si possano fare le corrispondenti operazioni con i segmenti. Per la somma e la sottrazione di due segmenti la cosa è evidente. Dando per scontato che sia possibile muoverli a proprio piacimento, basta collocarli uno dopo l’altro per la somma o soprapporli per la differenza. Per quanto riguarda il prodotto e la divisione si utilizza invece la nozione di proporzione. Infatti supponiamo di volere moltiplicare i segmenti d e c. Allora basta trovare una costruzione geometrica per cui valga una proporzione del tipo $1 : d = c : x$ in quanto, essendo il prodotto dei medi uguale al prodotto degli estremi, in tale caso il segmento x rappresenterà il prodotto di d per c. D'altra parte è ben noto come ottenere grandezze proporzionali in geometria: basta considerare triangoli simili.

Definizione 4.1. Il triangolo ABC si dice *simile* al triangolo $A'B'C'$ se l'angolo in A è uguale all'angolo in A', l'angolo in B è uguale all'angolo in B' e l'angolo in C è uguale all'angolo in C'.

Ovviamente la relazione di similitudine è una relazione di equivalenza, cioè è riflessiva, simmetrica e transitiva. Ricordando che nella geometria euclidea la somma degli angoli interni di un triangolo è un angolo piatto, è possibile dimostrare la seguente proposizione.

Proposizione 4.2. Dati due triangoli è sufficiente che due degli angoli siano uguali perché si possa asserire che sono simili. Dati due triangoli rettangoli, è sufficiente che uno degli angoli sia uguale per asserire che sono simili.

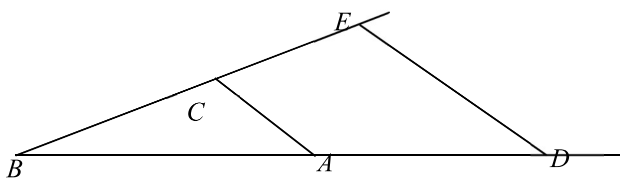
Il seguente teorema, di cui omettiamo la dimostrazione, gioca un ruolo fondamentale nella matematica greca.

Teorema 4.3. Due triangoli simili hanno i lati proporzionali. Più precisamente, supponiamo che A, B, C siano i vertici di un triangolo e A', B', C' i vertici di un altro triangolo. In tale caso se l'angolo in A è uguale all'angolo in A', l'angolo in B è uguale all'angolo in B', e l'angolo in C è uguale all'angolo in C', allora

$$AB : A'B' = AC : A'C' \quad \text{e} \quad AC : A'C' = BC : B'C'.$$

5. Calcolo dei segmenti in Cartesio

Utilizzando il teorema degli angoli simili Cartesio dice, con riferimento alla seguente figura,



. . . sia per esempio BA l'unità: se bisogna moltiplicare BD per BC devo soltanto aggiungere i punti A e C, poi tracciare DE parallela a CA, e BE è il risultato di questa moltiplicazione.

.

In altre parole si considerino due rette distinte per il punto *B* e due punti *D* e *C* su tali rette in modo che *BD* sia uguale a *d* e *BC* sia uguale a *c*. Sia inoltre *A* un punto della retta per *B* e *D* tale che *BA* sia unitario. Si tracci infine la parallela a *AC* per *D* e si denoti con *E* il punto di intersezione con la retta per *B* e *C*. Allora per la similitudine dei triangoli *CBA* e *EBD* risulterà che 1 : *BD* = *BC* : *BE*. In conclusione *BE* è il prodotto cercato.

Esercizio. Calcolare il prodotto di 3 per 5 in modo grafico (cioè utilizzando riga e compasso).

La stessa costruzione, dati *BE* ed *BD*, vale per la divisione.

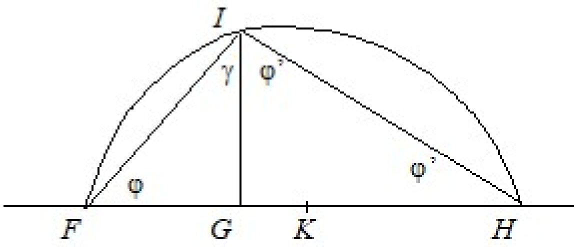
. . . Se invece bisogna dividere BE per BD, avendo unito i punti E e D, traccio AC parallela a DE e BC è il prodotto di questa divisione.

Cioè, se si deve dividere il segmento *d* per il segmento *c*, allora basta fissare un punto *E* in modo che *BE* sia uguale a *d* ed un punto *D* in modo che *BD* sia uguale a *c*. Tracciata allora la retta per *A* parallela ad *ED*, chiamo con *C* il punto di intersezione con la retta per *B* ed *E*. Il segmento *BC* è il risultato della divisione come si ricava dalla proporzione 1 : *BD* = *BC* : *BE*.

Esercizio. Calcolare 6/3 in modo grafico.

Anche nel caso di estrazione di radice quadrata abbiamo che il problema si traduce nello stabilire una opportuna proporzione.

. . . Se bisogna estrarre la radice quadrata di GH, aggiungo in



linea retta FG, che è l'unità, e dividendo FH in due parti uguali nel punto K, dal centro K traccio (la semicirconferenza) FIH, poi innalzando dal punto G una linea retta fino ad I ad angoli retti su FH, GI è la radice cercata.

Ad esempio se voglio trovare graficamente la radice 9 allora applico la seguente procedura:

1. Traccio un segmento *GH* di lunghezza 9
2. Prolungo a sinistra tale segmento con un segmento *FG* di lunghezza 1
3. Trovo il punto medio *K* del segmento *FH*
4. Traccio la circonferenza di centro *K* e diametro *FH* = 10
5. Alzo la perpendicolare dal punto *G*

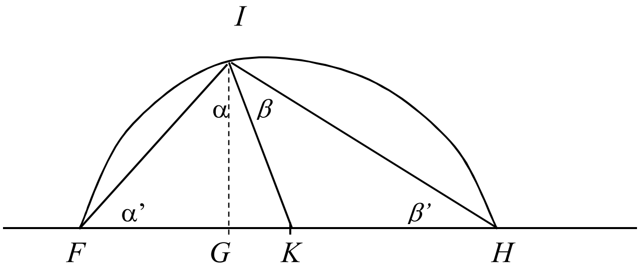
Il segmento *GI* misurerà esattamente 3, cioè la radice di 9.
Tale procedura è giustificata dai seguenti teoremi.

Teorema 5.1. (Secondo teorema di Euclide). L'altezza di un triangolo rettangolo è medio proporzionale delle proiezioni dei cateti.

Dim. Basta mostrare che i due triangoli rettangoli *FGI* e *IGH* sono simili. Per fare questo è sufficiente provare che $\varphi = \varphi'$. Ora, essendo la somma degli angoli di un triangolo un angolo piatto, $\varphi = 180-90-\gamma = 90-\gamma$ ed essendo il triangolo retto in *I*, risulta che $\varphi' = 90-\gamma$. Pertanto $\varphi = \varphi'$. □

Teorema 5.2. Un triangolo che abbia un lato *FH* uguale al diametro di una circonferenza ed il vertice opposto *I* sulla circonferenza è retto in *I*.

Dim. Tracciamo il segmento che congiunge il vertice I con il centro. Si ottengono due triangoli *FIK e IKH* che sono isosceli poiché due dei lati coincidono con i raggi. Allora l'angoli in I è uguale ad $\alpha+\beta=\alpha'+\beta'$ ma risulta anche uguale a 180-($\alpha'+\beta'$).



Da ciò si ricava che tale angolo è retto. □

Proposizione 5.3. La regola di Cartesio per l'estrazione della radice quadrata è corretta.

Dim. Poichè vale la proporzione

$$FG : GI = GI : GH.$$

ed essendo il prodotto dei medi è uguale al prodotto degli estremi, risulta

$$FG \cdot GH = GI^{2}$$

Pertanto, essendo $FG = 1$, $HG = GI^2$. □

In particolare è possibile estrarre la radice quadrata tramite il calcolo dei segmenti.

Esercizio. Trovare la radice di 7 in maniera grafica utilizzando cioè un righello ed un compasso.

6. Il *"Discorso sul Metodo"*

Il calcolo dei segmenti era alla base del metodo proposto da Cartesio, tuttavia si deve sottolineare che quello che Cartesio proponeva era una riduzione della geometria ai suoi elementi più semplici, i segmenti, e non una riduzione della geometria a manipolazione algebrica di numeri come avviene attualmente nella geometria analitica.[2]

[2] La geometria analitica intesa come completa riduzione al calcolo numerico non era possibile in quanto nel 1600 non era stata ancora data

Tutti i problemi della geometria si possono facilmente ridurre a tali termini che in seguito per costruirli basta conoscere la lunghezza di alcune rette.

Tali elementi semplici si possono manipolare con operazioni simili a quelle dell'aritmetica e pertanto è più corretto dire che con Cartesio si ha una algebrizzazione della geometria che pone la nozione di operazione alla base di tutto.

D'altra parte in Cartesio non vi era solo l'esigenza di ridurre la geometria a calcolo (di segmenti). Altrettanto importante era il processo inverso che consiste nella possibilità di interpretare ogni discorso algebrico in termini geometrici. In altre parole egli pensava si dovesse

- da un lato liberare la geometria dal ricorso obbligato alle figure che affaticavano inutilmente l'immaginazione
- da un altro lato dare significato alle operazioni dell'algebra per mezzo di una interpretazione geometrica.

Quanto poi all'analisi degli antichi e all'algebra dei moderni . . . , la prima è sempre siffattamente legata alla considerazione delle figure, che essa non può esercitare l'intelligenza senza affaticare di molto l'immaginazione; e, nella seconda, ci si è talmente assoggettati a certe regole e a certe cifre, che se ne è fatta un'arte confusa ed oscura, la quale tiene imbarazzato lo spirito, invece di (essere) una scienza che lo coltivi.

Scopo dichiarato di Cartesio è la ricerca di un metodo generale in contrasto con il modo frammentario con cui procedevano i greci antichi quando si trattava di trovare una dimostrazione o di risolvere un problema. Se infatti è certamente un merito dei greci il fatto che ogni dimostrazione sia rigorosamente controllabile nei sui singoli passaggi, niente viene detto da essi circa il metodo che si deve seguire per poter trovare nuovi teoremi e dimostrazioni. Pertanto restiamo disarmati di fronte ad ogni problema nuovo che si presenta e dobbiamo ogni volta procedere per tentativi.

Il metodo proposto da Cartesio per la geometria consisteva

---

una definizione di numero reale svincolata dall'intuizione del continuo (bisognerà aspettare per questo la fine del 1800).

- nell'indicare con lettere i dati e le incognite di un problema geometrico
- di tradurre le informazioni disponibili in equazioni
- nel semplificare, tramite calcoli algebrici, le equazioni quanto più possibile
- nel risolvere le equazioni risultanti da tale semplificazione in termini geometrici.

Pertanto abbiamo un passaggio del tipo

Geometria $\rightarrow$ Algebra $\rightarrow$ Geometria

piuttosto che un annullamento della geometria. Ad esempio dopo aver tradotto un problema geometrico in una equazione di secondo grado era opportuno semplificare al massimo tale equazione. Giunti però alla forma più semplice possibile la risoluzione della equazione finale doveva essere di tipo grafico. Pertanto la risoluzione grafica (detta 'costruzione') di semplici equazioni di secondo grado, in particolare il calcolo grafico di una radice quadrata, era un elemento essenziale della teoria di Cartesio.
Ecco che cosa dice Cartesio nella sua *Geometria*

> *Così, volendo risolvere qualsiasi problema, si deve innanzi tutto considerarlo come risolto, e si devono dare dei nomi a tutte le linee che sembrano necessarie per la sua costruzione, sia quelle ignote, sia alle altre. Poi, senza fare alcuna differenza tra queste linee, note ed ignote, bisogna affrontare le difficoltà secondo l'ordine che mostra nella maniera più naturale in che modo tali linee siano in rapporto tra loro, fino a che non si sia trovato modo di esprimere una medesima quantità in due maniere diverse: ciò si chiama un'equazione* (*in una sola incognita*) *poiché i termini di una di queste due espressioni sono uguali a quelli dell'altra.*

Si noti che Cartesio tratta prevalentemente problemi che si traducono in una equazione ad una sola incognita e che l'idea di luogo geometrico, insieme dei punti le cui coordinate verificano una equazione a due variabili, non è presente nella sua opera se non in modo saltuario.

Concludiamo questo paragrafo sottolineando che le teorie matematiche di Cartesio erano strettamente legate al suo sistema

filosofico più generale. Basti pensare che il suo libro La Geometria non venne pubblicato come un trattato a sé stante ma come una delle tre appendici del *"Discorso sul metodo"* il cui titolo completo è
"Discorso sul metodo per ben condurre la propria ragione e cercare la verità nelle scienze"
e che tali appendici avevano appunto il ruolo di illustrare il suo metodo filosofico generale. I precetti fondamentali di tale metodo erano:

- Il precetto dell'evidenza;
- Il precetto dell'analisi;
- Il precetto della sintesi;
- Il precetto del computo completo.

Ed il primo era, di non accettare cosa alcuna per vera quando non la riconoscessi evidentemente per tale: cioè, di evitare studiatamente la precipitazione e la prevenzione; e di non accogliere nei miei giudizi nulla di più di ciò che si presentasse sì chiaramente e sì distintamente al mio spirito da non poter aver motivo alcuno di metterlo in dubbio.

Il secondo, di dividere ogni difficoltà, ch'io esaminassi, in parti elementari fino al limite del possibile e quanto sarebbe richiesto per trovarne la migliore soluzione.

Il terzo, di condurre per ordine i miei pensieri, cominciando dagli oggetti più semplici e più facili da conoscer, per salire a poco a poco e come per gradi alla conoscenza dei più complessi . . .
E l'ultimo, di fare, in ogni argomento, enumerazioni così complete e verifiche così generali da essere sicuro di nulla omettere.

7. Aritmetizzazione della geometria: la sparizione delle figure

"Il lettore non troverà figure in questo lavoro. I metodi che esporrò non richiedono costruzioni né geometriche né meccaniche,

ma solamente operazioni algebriche, soggette a una procedura regolare e uniforme.”
Jean Louis Lagrange*, Mécanique Analytique*)

Se si considera l'importanza che la geometria di Euclide aveva avuto nella cultura dei greci ed in quella successiva, ci si rende conto di quanto fosse un evento rivoluzionario e sorprendente la scoperta delle geometrie non euclidee. I pensatori precedenti avevano costantemente ritenuto che vi fosse una sola geometria vera e che le sue leggi fossero necessariamente quelle di Euclide. Inoltre il modo di procedere geometrico era sempre stato visto come un modello a cui ispirarsi in tutti gli altri campi del sapere. L'apparire di tali nuove geometrie confutava queste convinzioni perché se più teorie dello spazio contrastanti tra loro sono logicamente possibili e se solo una di queste poteva essere vera, allora la geometria, e più in generale la matematica non poteva più essere considerata lo strumento per giungere alla verità.

D’altra parte il mettere in discussione il ruolo centrale ed assoluto della geometria faceva nascere in modo sempre più pressante l'esigenza di trovare una nuova base alla matematica. Si deve anche tenere conto che era necessario inquadrare anche la nuova matematica nata dai metodi infinitari del calcolo differenziale.

I passi di una tale nuova fondazione della matematica consistirono essenzialmente:

a) nell'aritmetizzazione della geometria e dell'analisi,

b) nella teoria degli insiemi di G. Cantor.

Abbiamo già visto che con Cartesio la geometria era stata ridotta ad un calcolo dei segmenti e quindi, in un certo senso, ridotta all'algebra. Per poter fare completamente a meno della geometria era allora necessario un ulteriore passo: sostituire al calcolo dei segmenti un calcolo numerico che si fondasse su di una definizione di numero reale completamente indipendente dall'intuizione geometrica.

Nel prossimo capitolo mostreremo come ciò sia possibile definendo prima i numeri naturali, poi gli interi relativi, poi i razionali ed infine i reali. In questo paragrafo daremo per scontata la conoscenza dei numeri reali e mostreremo come tali numeri siano sufficienti a definire le nozioni geometriche. Infatti, come è noto, con lo sviluppo della geometria analitica, la geometria diverrà un capitolo dell'algebra lineare sul campo dei numeri reali.

Teorema 7.1. Sia R l'insieme dei numeri reali, chiamiamo *piano euclideo* il prodotto cartesiano $R\times R$, e *punti* i suoi elementi. Inoltre chiamiamo *retta* un insieme di punti che verifichi una equazione lineare del tipo $ax+by+c = 0$. Allora la struttura ottenuta in tale modo verifica tutti gli assiomi della geometria euclidea.

Dim. Proviamo ad esempio che per due punti distinti $(\underline{x}_0,\underline{y}_0)$ e $(\underline{x}_1,\underline{y}_1)$ passa una ed una sola retta. Tale problema si traduce in quello di trovare a, b, c (non tutti nulli) tali che

$$a\underline{x}_0+b\underline{y}_0 +c = 0$$
$$a\underline{x}_1+b\underline{y}_1 +c = 0$$

Si tratta di un sistema omogeneo di due equazioni nelle incognite a, b e c e dalla teoria dei sistemi lineari si sa che se i due punti sono diversi tra loro tale sistema ammette infinite soluzioni e che due diverse soluzioni sono proporzionali tra loro. Ne segue che tutte queste soluzioni rappresentano una stessa retta e ciò prova l'esistenza e l'unicità della retta per $(\underline{x}_0, \underline{y}_0)$ e $(\underline{x}_1, \underline{y}_1)$. Più in generale, possiamo dire che il passaggio per un punto è una condizione lineare omogenea e che quindi la retta che passa per due punti prefissati distinti si trova imponendo due condizioni lineari omogenee. Da ciò segue l'esistenza e l'unicità di tale retta.

In modo analogo si prova il quinto postulato. Si tratta di verificare che dato un punto $P = (\underline{x},\underline{y})$ ed una retta r, che supponiamo di equazione $ax+by+c = 0$, allora esiste una ed una sola retta passante per $(\underline{x},\underline{y})$ e parallela ad r. Ma anche il parallelismo è una condizione lineare omogenea poiché una retta r' con coefficienti a' e b' è parallela ad r se e solo se $ab'-a'b = 0$. Pertanto esiste ed è unica la retta per P parallela ad r. I rimanenti assiomi si dimostrano in modo altrettanto semplice. □

Un processo di riduzione al calcolo dei numeri reali è stato poi fatto anche per quanto riguarda l'analisi matematica. Infatti nascono le attuali definizioni di limite, derivata, funzione continua che fino alla prima metà dell'ottocento si basavano sulla intuizione del continuo geometrico. In definitiva si attua quello che viene a volte chiamato *"processo di aritmetizzazione della matematica"* in cui tutto viene ridotto ai numeri reali. Poiché i numeri reali si possono definire a partire dagli interi, riappare l'antica idea della scuola Pitagorica per cui tutto è riconducibile ai numeri interi.

Nel prossimo capitolo vedremo come sia possibile costruire la struttura algebrica dei numeri reali.

8. Intuizione geometrica e falsi teoremi euclidei

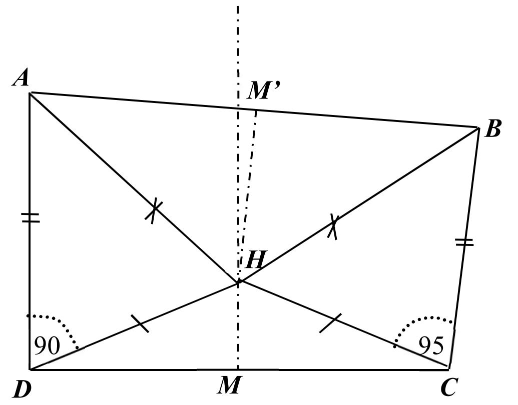
Le dimostrazioni presenti nei libri di Euclide sono sempre molto belle ed intuitive. Infatti in esse sono sempre presenti sia il rigore logico della deduzione sia l'interpretazione intuitiva dei singoli passi di tale deduzione. Tuttavia a volte l'intuizione geometrica trae in inganno (come abbiamo già visto nel caso della equiscomponibilità) ed in questo paragrafo mostro alcuni esempi. Nel seguito riporto la dimostrazione, di tipo euclideo, del fatto che 5 = 0. Non so chi l'abbia inventata in quanto mi è stata raccontata da un collega a cui è stata raccontata da un altro collega . . . Lascio a chi legge il compito non facile di trovare dove è l'errore.

Teorema 8.1. Il numero 5 è uguale al numero 0.

Dim. Tracciamo un segmento *DC* ed alziamo da *D* un segmento *AD* perpendicolare a *DC.* Alziamo da *C* un segmento di uguale lunghezza che faccia un angolo di 95 gradi col segmento *CD.* Otteniamo un quadrilatero di vertici *A*, *B*, *C*, *D.* Tracciamo ora l'asse di *DC* (dal punto medio *M* di *DC*) e l'asse di *AB* (dal punto medio *M'* di *AB*). Poiché *DC* non è parallelo ad *AB* i due assi non sono paralleli tra di loro e pertanto si incontrano in un punto *H.* Si vengono pertanto a formare due triangoli *AHD* e *BHC* che risultano uguali. Infatti *AD* è uguale a *BC* per costruzione, *AH* = *HB* perché il triangolo *AHB* è isoscele, *DH* = *CH* perché il triangolo *DHC* è isoscele. Dal fatto che *AHD* sia uguale a *BHC* comporta che l'angolo $A\hat{D}H$ sia uguale all'angolo $H\hat{C}B$. Essendo il triangolo *DHC* isoscele, risulta anche che $H\hat{D}C = D\hat{C}H$. In definitiva possiamo concludere che

$$A\hat{D}C = A\hat{D}H + H\hat{D}C = H\hat{C}B + H\hat{D}C = D\hat{C}B$$

e quindi che 90 = 95. Sottraendo 90 si ottiene che 0 = 5.



□

Corollario 8.2. Io sono l'uomo più bello, più intelligente, più simpatico del mondo, inoltre sono anche il più grande matematico che sia mai esistito.

Dim. Consideriamo l'insieme costituito da me, dall'uomo più bello, dall'uomo più intelligente, dall'uomo più simpatico e dal più grande matematico che sia mai esistito arriviamo ad un insieme X con 5 elementi. Poiché abbiamo dimostrato che $5 = 1$, X ha un solo elemento e quindi tutte le persone che ho elencato sono in realtà una unica persona. Ciò prova il corollario[3]. □

Un altro falso teorema (in cui viene fatto un errore simile a quello del teorema ora esposto) è il seguente.

Teorema 8.3. Tutti i triangoli sono isosceli.

---

[3] Questa dimostrazione è un ovvio adattamento di una dimostrazione che si racconta abbia fatto Bertrand Russell. Pare che a Russell sia stato chiesto come possa essere accettata una cosa tanto strana per cui a partire da una asserzione falsa possa essere dimostrata qualsiasi altra asserzione e che per sfida gli sia stato chiesto come da 2=1 si possa dimostrare, ad esempio, che Russell è Dio. La risposta di Russell fu appunto che essendo Russell e Dio due cose distinte ed essendo 2 =1, allora Russell non poteva che essere uguale a Dio.

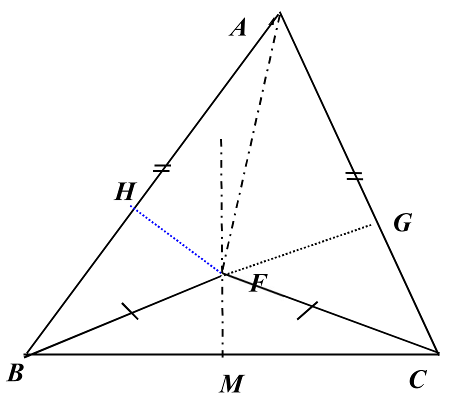
Dim. Consideriamo un triangolo di vertici *A*, *B*, *C* ed indichiamo con *M* il punto medio del segmento *BC*. Alziamo da *M* la perpendicolare a *BC* e da *A* la bisettrice dell'angolo in *A*. Due sono i casi: che le due rette si incontrino in un punto o che siano parallele..

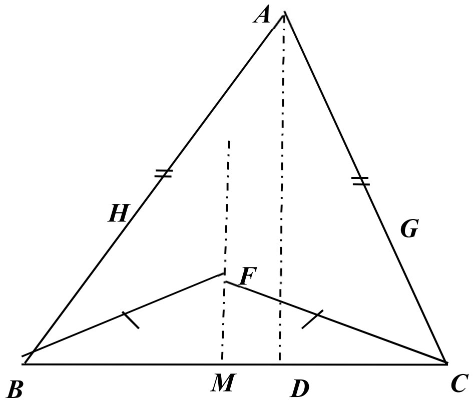
Nel primo caso (prima figura affianco) indichiamo con *F* il punto di incontro e, dopo avere tracciato i segmenti *FB* ed *FC*, a partire da *F* tracciamo le perpendicolari *FG* e *FH* ad *AC* e *AB*. I due triangoli *AHF* e *AFG* sono uguali in quanto sono rettangoli, hanno un lato in comune e, essendo la retta *AF* la bisettrice, l'angolo *HAF* è uguale all'angolo *FAG*. D'altra parte i due triangoli *BFM* e *MFC* sono uguali in quanto rettangoli con *BM* = *MC* ed il lato *MF* in comune. Ne segue che essendo *BF* = *FC* e *HF* = *FG* i due triangoli rettangoli *BHF* e *FCG* sono uguali. Concludendo

$$AB = BH+HA = CG+AG = AC$$

e quindi *ABC è isoscele.*

Consideriamo il caso in cui la bisettrice e l'asse *BC* non si incontrano (seconda figura affianco). Allora le due rette sono parallele e quindi la bisettrice risulta perpendicolare a *BC*. Ne segue che i due triangoli rettangoli *BDA* e *ADC* avendo un lato in comune ed i due angoli in *A* uguali sono uguali. Pertanto *BA* = *AC*.



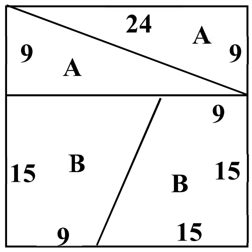


□

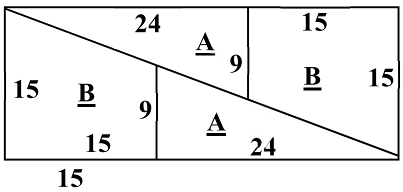
Ora mostro un esempio di coppia di figure che sembrerebbero equiscomponibili ma non lo sono.

Proposizione 8.4. (Paradosso della scomposizione del quadrato). E' possibile scomporre un quadrato di area 576 ed ottenere un rettangolo di area 585.

Dim. Costruiamo un quadrato di lato uguale a 24 e tagliamolo in due rettangoli di lati 24 e 9 e 24 e 15.



AREA = 24×24
= 576



AREA = (15+24)×15
=39×15 = 585

Tagliamo il rettangolo piccolo in due triangoli rettangoli di cateti 9 e 24 ed il rettangolo grande in due trapezi rettangoli di basi 15 e 9 e di altezza uguale a 15. Ricomponiamo poi tali pezzi in modo da ottenere il rettangolo dato in figura di lati 15 e 39 (si vedano le figure). Per costruzione il quadrato ed il rettangolo sono equi-scomponibili. Inoltre l'area del quadrato è uguale a $24\times24 = 576$, l'area del rettangolo risulta uguale a $15\times39 = 585$.[4] □

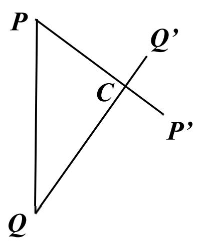
9. Altri falsi teoremi

Di natura completamente differente è il seguente teorema, dovuto a Proclo, che nega il postulato delle parallele. Ricordiamo che tale postulato viene enunciato dicendo che:

> *supponiamo che una retta cada su due rette formando due angoli interni e dalla stessa parte minori di due retti, allora le due rette prolungate illimitatamente verranno ad incontrarsi dalla parte in cui sono gli angoli minori di due retti.*

Sembra invece possibile dimostrare il seguente "falso" teorema[5].

Teorema 9.1. Supponiamo che una retta r cada su due rette s e t formando due angoli interni e dalla stessa parte minori di due retti, allora è possibile prolungare illimitatamente le due rette da quella parte in cui sono gli angoli minori di due retti senza che si incontrino mai.



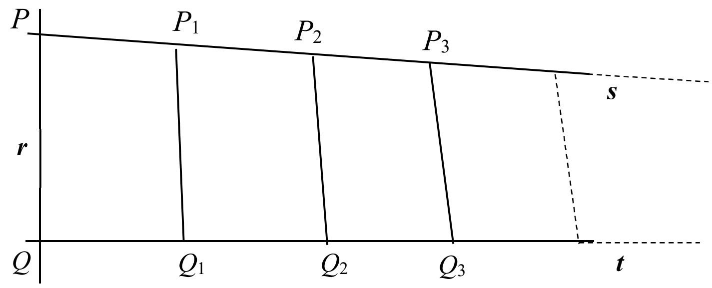
Dim. Per prima cosa osserviamo che dato un segmento PQ se da P e da Q partono due segmenti PP' e QQ', tali che $PP'+QQ' \leq PQ$ allora tali segmenti non si possono incontrare. Infatti se si incontrassero in un punto C, come in figura, allora per la diseguaglianza triangolare, si avrebbe l'assurdo per cui

$$PQ < PC+CQ \leq PP'+QQ' \leq PQ.$$

---

[4] Chi non fosse convinto dei disegni può provare a fare esperimenti con un foglio di carta quadrettato ed un paio di forbici.

[5] Il teorema più che essere falso, suggerisce l'esigenza di precisare cosa si debba intendere per "illimitatamente".



Passando alla dimostrazione, siano P e Q i due punti di intersezione di r con le rette s e t ed indichiamo con P_1 e Q_1 due punti su s e t tali che i segmenti PP_1 e QQ_1 siano entrambi uguali alla metà di PQ. Allora, per quanto ora detto, questi due segmenti non si possono incontrare. Prolunghiamo ora i segmenti PP_1 e QQ_1 in due segmenti P_1P_2 e Q_1Q_2 che siano entrambi uguali alla metà di P_1Q_1. Per lo stesso motivo detto prima questi due segmenti non si possono incontrare. Avanziamo ulteriormente percorrendo i segmenti P_2P_3 e Q_2Q_3 entrambi di lunghezza pari alla metà di P_2Q_2. Tali segmenti non si possono incontrare ... Ripetendo all'infinito tale ragionamento si ottiene che nessuno dei prolungamenti costruiti su s e t in questo modo si potrà mai incontrare.[6]

□

Un altro modo di negare il postulato delle parallele è provare che esiste un triangolo la cui somma degli angoli interni è maggiore di un angolo piatto come viene fatto nella seguente proposizione.

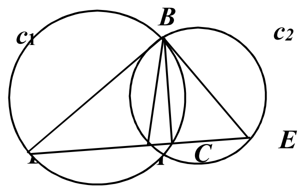
Proposizione 9.2. Esiste un triangolo la cui somma degli angoli interni è maggiore di un angolo piatto.

Dim. Consideriamo due cerchi c_1 e c_2 che si incontrano nel punto B e tracciamo da tale punto i due diametri BD e BE. Indichiamo inoltre con A e C i due punti di intersezione della retta DE con

[6] Questa falsa dimostrazione utilizza un argomento simile a quello del paradosso di Achille e la tartaruga anche se quest'ultimo è più intrigante in quanto può essere visto sia rispetto alla variabile tempo (in nessun istante Achille supera la tartaruga) che alla variabile spazio (in nessun punto Achille supera la tartaruga).

le circonferenze. L'angolo in C del triangolo BCE è retto perché è un angolo alla circonferenza c_1 relativo al diametro BD.

L'angolo in A del triangolo BAE è retto perché è angolo alla circonferenza c_2 relativo al diametro BE. In conclusione il triangolo ABC ha due angoli retti e quindi la somma dei suoi angoli interni è maggiore di un angolo piatto. □



Nello studio delle circonferenze si ricorre spesso all'intuizione per cui una circonferenza è in un certo senso il limite dei poligoni regolari di n lati inscritti nella circonferenza stessa (per n che tende all'infinito). Questa idea è applicata ad esempio per calcolare un valore approssimato della lunghezza di una circonferenza. Oppure è utilizzata per dimostrare che l'area di una circonferenza è uguale alla lunghezza della circonferenza per la metà del raggio.

Proposizione 9.3. L'area A di un cerchio è uguale alla lunghezza della circonferenza per il raggio diviso due. In breve:

$$A = \pi \cdot r^2.$$

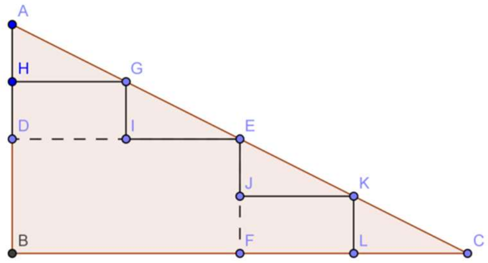
Dim. Abbiamo gia provato che l'area di un poligono regolare è uguale al perimetro per l'apotema diviso due. Per ogni $n \geq 3$ indichiamo con A_n l'area racchiusa in un poligono regolare P_n di n lati inscritto nel cerchio, e con a_n e p_n rispettivamente la relativa apotema ed il relativo perimetro. Allora $A_n = a_n \cdot p_n/2$. Per n che tende all'infinito A_n tende all'area del cerchio, a_n al raggio e p_n alla lunghezza della circonferenza. Pertanto la proposizione è dimostrata.

Da notare che è possibile anche considerare i poligoni regolari circoscritti al cerchio. In tale caso, se indichiamo con $\underline{P}_n$ uno di tali poligoni, con $\underline{p}_n$ il suo perimetro e con $\underline{A}_n$ la sua area, risulta che $\underline{A}_n = r \cdot \underline{p}_n/2$. Ancora una volta passando a limite abbiamo una dimostrazione della proposizione. □

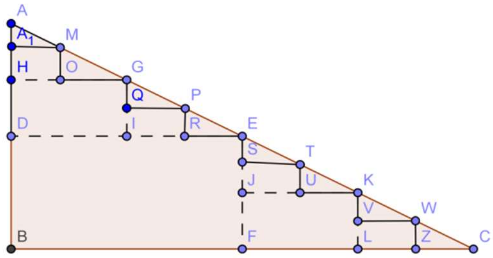
Purtroppo anche se l'enunciato del teorema è corretto, la sua dimostrazione non lo è. Per rendercene conto applichiamo lo stesso metodo per dimostrare una proposizione falsa.

Proposizione 9.4. Dato un triangolo rettangolo la lunghezza dell'ipotenusa è uguale alla somma delle lunghezze dei cateti. Pertanto il teorema di Pitagora (che asserisce che è uguale alla radice della somma dei quadrati di tali lunghezze) è falso.

Dim. Dato un triangolo rettangolo di vertici A, B e C consideriamo la spezzata $ADEFC$ che si ottiene considerando il punto medio F del cateto BC ed il punto medio D del cateto AB. Tale



spezzata ha
ovviamente lunghezza $AD+DE+EF+FC = AD+EF+DE+FC = AD+DB+BF+FC = AB+BC$. Ripetiamo tale operazione per ciascuno dei due triangoli ADE e EFG ottenendo la spezzata $AHGIEJKLC$ la quale, ovviamente, ha ancora lunghezza $AB+BC$. Applichiamo tale procedura a ciascuno dei quattro triangoli AHG, GIE, EJK, KLC che si sono formati in questo modo, ottenendo la spezzata "più vicina all'ipotenusa" $AA1MOGQPRESTUKVWZC$ e così via (si veda la figura che segue).



Si ottiene in questo modo una successione di spezzate tutte di lunghezza $AB+BC$ che risultano man mano sempre più vicine all'ipotenusa. Essendo l'ipotenusa il limite di tale successione di spezzate, risulta che la sua misura deve essere $AB+BC$.

Da notare che avremmo pure potuto costruire una successione di spezzate "al di sopra" dell'ipotenusa e ripetere tale tipo di argomentazione. □

10. Divagazioni: figure geometriche ed equiscomponibilità

Che è quel ch'adunque divide l'aria dall'acqua?» È aria o acqua? (Leonardo da Vinci: *Codice Atlantico*, UTET 1966: 546)

In questo paragrafo analizziamo la questione su come la moderna geometria analitica possa rappresentare gli enti geometrici proposti dai greci. Infatti, riferendoci alla geometria piana, è chiaro come essa rappresenta i punti, le rette e le circonferenze ma non è altrettanto chiaro come essa rappresenta figure geometriche come il cerchio, il quadrato ed il triangolo. Tali figure sono insiemi aperti o insiemi chiusi ? Da tale questione dipende ovviamente anche il significato da assegnare alle varie nozioni della geometria di Euclide quando si passa all'ambito analitico. Consideriamo ad esempio la nozione di equiscomponibilità che è del tipo "foglio di carta + forbici + spostamenti" ed appare molto naturale. Abbiamo già visto che è possibile descrivere tale nozione in termini di partizioni e di isometrie.

Ma la definizione proposta non è completamente convincente. Ad esempio, possiamo affermare che l'esagono del teorema 9.7 del Capitolo 1 si scompone nei triangoli A_1, A_2, B_1,...,B_5, solo se tali triangoli fossero a due a due disgiunti. Per ottenere questo dovremmo precisare per ogni punto di contatto tra due triangoli vicini se tale punto appartiene ad un triangolo oppure all'altro. Cosa questa possibile ma certamente poco naturale. Si torna allora alla questione da cui siamo partiti: *che cosa è un triangolo?* Se si accetta il fatto che tre punti non allineati individuano un solo triangolo, bisogna stabilire se tale triangolo contiene tutti i punti dei lati (è un insieme chiuso), nessun punto dei lati (è aperto) oppure altro. Assumiamo ad esempio che per figura geometrica, e quindi in particolare per triangolo, si intenda un insieme chiuso.

Definizione 10.1. Chiamiamo "*figura geometrica*" un sottoinsieme chiuso del piano.

In tale caso vediamo come si potrebbe la nozione di scomposizione. E' ragionevole ad esempio richiedere che l'intersezione $X_i \cap X_j$ pur non essendo vuota sia "trascurabile" e che trascurabile significhi ad esempio "di misura nulla".

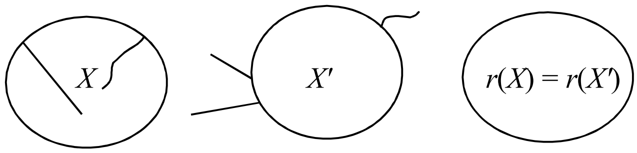
Definizione 10.2. Diciamo che le figure geometriche $X_1,\dots,X_n$ sono una *scomposizione* di una figura geometrica X se

- $X = X_1 \cup \dots \cup X_n$
- $X_h \cap X_k$ è un insieme di misura nulla per ogni $i \neq j$.

Una ulteriore definizione di scomposizione si ottiene sostituendo la seconda condizione con la richiesta che l'interno di X_h sia disgiunto dall'interno di X_k. Queste definizioni sembrano funzionare, tuttavia coinvolgono le nozioni di misura o di interno in modo un po' artificiale se si pensa al significato intuitivo di "tagliare a pezzi" su cui si basa l'idea di scomposizione. In realtà la nostra intuizione è che di fatto non ci interessano i punti della frontiera poiché, in un certo senso senso questa è una questione che riguarda l'analisi matematica e non la geometria. Un modello matematico che giustifichi questo convincimento è il seguente.

Definizione 10.3. Chiamiamo *regolarizzatore* l'operatore r che si ottiene ponendo $r(X) = c(i(X))$[7]. Chiamiamo *chiuso regolare*, ogni punto fisso di r.

In altre parole, $r(X)$ è l'insieme di punti che si possono ottenere come limite di una successione di elementi interni ad X. L'effetto del regolarizzatore è quello di togliere da un insieme di punti *"pezzi"* o *"tagli"* di dimensione inferiore alla dimensione dello spazio ambiente. Nel seguito sono riportati esempi di due insiemi X ed X' che si regolarizzano entrambi nello stesso cerchio (nella prima figura le linee sono tagli):



[7] Indichiamo con i e c gli operatori di interno e di chiusura..

Si dimostra che tale operatore è monotono ed idempotente e ciò comporta che i chiusi regolari sono tutti e soli gli insiemi che sono il risultato del processo di regolarizzazione.

Definizione 10.4. Chiamiamo *figura geometrica* ogni chiuso regolare.

Un cerchio chiuso o un triangolo chiuso sono figure geometriche mentre una linea o un numero finito di punti non lo sono. La classe dei chiusi regolari dello spazio euclideo è una algebra di Boole completa (senza atomi). Le due operazioni reticolari $\vee$ e $\wedge$ sono definite ponendo $X\vee Y = X\cup Y$ e $X\wedge Y = r(X\cap Y)$. Pertanto se X ed Y sono due triangoli con un lato o un vertice in comune allora risulta che $X\wedge Y = \emptyset$ a dispetto del fatto che $X\cap Y \neq \emptyset$. Ciò suggerisce di definire la scomposizione come una partizione nell'algebra di Boole degli insiemi regolari.

Definizione 10.5. Chiamiamo *scomposizione* di una figura geometrica X una sequenza $X_1,\dots,X_n$ di figure geometriche tali che
- $X = X_1\vee\dots\vee X_n$
- $X_i\wedge X_j = \emptyset$ per ogni $i\neq j$.

Le scomposizioni che abbiamo fino ad ora visto rientrano nella definizione che ora abbiamo dato.

CAPITOLO 3
DAI NUMERI AL CONTINUO GEOMETRICO

Dio creò i numeri naturali, tutto il resto è opera dell'uomo.
Leopold Kronecker (1823-1891)

1. Un punto di partenza: i numeri naturali (terne di Peano)

Attualmente una buona parte della matematica si basa sulla nozione di numero, in particolar modo di numero reale. La geometria di Euclide con il relativo metodo deduttivo viene sostituita in larga parte dall'algebra lineare che, oltre ad essere un potente strumento per afrontare vari problemi, è anche la base migliore su quale si sviluppa il calcolo differenziale. Nelle università un punto viene definito come una coppia di numeri reali, una retta come l'insieme dei punti che soddisfano un'equazione lineare e così via.

Questo comporta che i fondamenti della matematica si riconducono ai fondamenti alla teoria dei numeri reali. Come vedremo in questo capitolo questi fondamenti si realizzano tramite una seccessione di costruzioni che partono dai numeri naturali ed arrivano alla costruzione dei reali che poi permette di fare il resto (in particolare la geometria).

naturali → relativi → razionali → reali → geometria →...[1]

Per quanto riguarda il primo passo, cioè la definizione dei numeri naturali, viene utilizzato il metodo assiomatico utilizzando il sistema di assiomi proposti da Dedekind e Peano.[2] Tali assiomi si basano sull'idea per cui l'insieme dei numeri naturali è il frutto del processo di aggiungere un nuovo elemento ad un elemento dato (si veda il raccontino di Zavattini alla fine del capitolo). D'altra parte se esaminiamo i metodi utilizzati in passato per "contare le pecore del proprio gregge" o "i giorni che passano" ci accorgiamo che essi consistono nel mettere, ad esempio, tacche su di un pezzo di legno e nell'aggiungere una tacca per ogni pecora rientrata nel recinto o per ogni giorno che passa. Altro

---

[1] Le costruzioni utilizzano anche alcune nozioni elementari di teoria degli insiemi ed in particolare, come vedremo, utilizzano l'accettazione dell'infinito attuale, in contrasto con le idee di Aristotele.

[2] Tali numeri costituiscono una nozione tanto immediata che probabilmente pretendere di definirla, quindi di ridurla a termini più semplici, non ha molto senso. Tuttavia lo sforzo di definirli ha il vantaggio di mettere in rilievo alcune delle proprietà essenziali di cui godono.

metodo può essere stato il mettere sassolini in un recipiente e nell'attività di aggiungere un sassolino se necessario.

Ecco la principale definizione.

Definizione 1.1. Chiamiamo *terna di Peano* una struttura algebrica (S, s, z_0), in cui s è un'operazione 1-aria e $z_0 \in S$ è un elemento designato, che verifica i seguenti assiomi:

P1 $s : S \to S$ è una funzione iniettiva

P2 $z_0 \notin s(S)$, cioè z_0 non è il successivo di nessun elemento

P3 $z_0 \in D$ e $s(D) \subseteq D \Rightarrow D = S$ (per ogni $D \subseteq S$).

La funzione $s : S \to S$ viene chiamata *funzione successore*, l'elemento z_0 viene chiamato *elemento nullo*, $s(x)$ il *successivo di x*. L'assioma ***P3*** è quello più importante e prende il nome di *principio di induzione matematica*.[3] Un modo diverso di scriverlo è il seguente:

$$((z_0 \in X) \text{ e } (x \in X \Rightarrow s(x) \in X)) \Rightarrow X = S.$$

Quando si propone una teoria si deve anche provare che ne esiste almeno un modello: in caso contrario la teoria parlerebbe del nulla. Pertanto dovremmo provare che esiste almeno una terna di Peano, cosa che ha a che fare con la questione dell'infinito e che faremo in seguito

E' anche interessante far vedere che molte strutture che i matematici utilizzano usualmente non sono terne di Peano: ecco alcuni esempi.[4]

Sia Z l'insieme degli interi relativi e definiamo s ponendo $s(x) = x+1$, allora $(Z, s, 0)$ non è una terna di Peano. Infatti l'assioma *P2* non vale in quanto 0 è il successore di -1. Inoltre non vale

[3] Da notare che tale teoria è espressa "al secondo ordine". Ciò significa che si utilizza un linguaggio in cui si applica un quantificatore ("per ogni") a sottoinsiemi D di S. In logica matematica, come vedremo, normalmente si considerano invece teorie "del primo ordine" in cui è possibile quantificare solo su elementi di S. Esiste comunque anche una teoria del primo ordine delle terne di Peano che vedremo nel seguito.

[4] Dare tali esempi sarebbe scorretto da un punto di vista metodologico. Infatti se stiamo "fondando" la matematica non possiamo riferirci ad esempi che attingono da una matematica non ancora fondata. La scorrettezza è solo apparente in quanto il loro ruolo è didattico (attingere ad una intuizione già esistente) e non matematico. D'altra parte che senso avrebbe l'impresa di "fondare la matematica" se non si avesse in mente già una idea dell'oggetto-matematica da fondare?

nemmeno *P3*. Infatti l'insieme D degli interi maggiori o uguali a zero pur verificando le due condizioni $0 \in D$ e $D \subseteq s(D)$ non coincide con Z.

Sia R^+ l'insieme dei reali maggiori o uguali a zero, allora (R^+, s, 0) non è una terna di Peano. Infatti anche se gli assiomi *P1* e *P2* sono soddisfatti *P3* non è soddisfatto.

Sia Z/m l'insieme degli interi modulo m e consideriamo la struttura (Z/m, s, [0]) dove $s([x]) = [x]+[1] = [x+1]$. E' evidente che *P3* è verificata, invece *P2* non vale in quanto [0] è successore di $[m\text{-}1]$. Questo mostra che (Z/m, s, [0]) non è una terna di Peano.

Problema: Esiste una terna di Peano con 5 elementi ?
Problema. Dimostrare che (R^+, s, 0) non è una terna di Peano fornendo almeno due esempi di insieme per cui non vale il principio di induzione.
Problema. Consideriamo la struttura (N_0, s, z_0) con N_0 insieme dei numeri naturali, $z_0 = 0$ ed s definita dal porre $s(n) = 2n+1$. Dire se tale struttura è una terna di Peano.

Esercizio. Consideriamo la struttura (D, s, z_0) con D insieme dei numeri naturali dispari, $z_0 = 0$ ed s definita dal porre $s(n) = n+2$. Dire se tale struttura è una terna di Peano.
Esercizio. Consideriamo la struttura (S, s, z_0) con S insieme dei numeri naturali maggiori o uguali a 5, $z_0 = 5$ e porre $s(n) = n+1$. Dire se tale struttura è una terna di Peano.
Esercizio. Sappiamo che il prodotto diretto di due gruppi è un gruppo e lo stesso si può dire per gli anelli o per i reticoli. Consideriamo il prodotto diretto di una terna di Peano (P, s, z_0) per se stessa, cioè la struttura ($P\times P$, s, (z_0,z_0)) definita ponendo $s((x,y)) = (s(x),s(y))$. Dire se tale struttura è ancora una terna di Peano.

2. Principio di induzione

In una terna di Peano è possibile effettuare due cose di particolare importanza: le dimostrazioni tramite il principio di induzione e le definizioni per ricorsione.

Proposizione 2.1 (Principio di induzione[5]). Supponiamo che una proprietà P sia definita in una terna di Peano (S,s,z_0) e che P verifichi le due seguenti affermazioni:

a) P è verificata da z_0

b) se P è verificata da x allora è verificata da $s(x)$,

allora è possibile concludere che

c) P è verificata per ogni $x \in S$.

Dim. Sia D l'insieme degli elementi che verificano P, allora D contiene z_0 ed è tale che $s(D) \subseteq D$. Pertanto tale insieme coincide con S e ciò prova che P è verificata per ogni $x \in S$. □

Un modo più compatto per rappresentare il principio di induzione è il seguente dove, in accordo con l'uso dei logici, scriveremo $P(x)$ per dire che x verifica la proprietà P

se *a*) $P(z_0)$

e *b*) $\forall x(P(x) \Rightarrow P(s(x)))$,

allora *c*) $\forall x\, P(x)$.

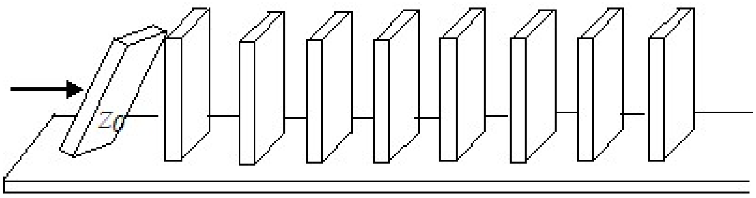
Il principio di induzione può essere visualizzato al modo seguente. Consideriamo la seguente figura in cui i pezzi del gioco domino sono poggiati su di un tavolo (infinito) uno dopo l'altro ed indichiamo il primo pezzo della fila con z_0. Vale la regola che se un pezzo cade (a destra) allora il pezzo successivo cade

$$\forall x Cade(x) \rightarrow Cade(s(x)).$$

poi supponiamo che valga

$$Cade(z_0)$$

Allora vale $\forall x Cade(x)$, cioè tutti i pezzi cadono.



[5] Non bisogna confondere tale principio, che appartiene alla matematica, con il principio di induzione in fisica. In fisica il principio di induzione è quello che permette di passare da una serie di esperimenti alla formulazione di una teoria. Ad esempio, poiché in tutte le nostre esperienze passate un corpo libero cade verso la terra, possiamo formulare la teoria "per ogni corpo x, se x è libero allora x cadrà verso la terra".

Esempio. Supponiamo che, nell'ambito del campo dei numeri reali, si voglia dimostrare che per ogni $n \in N$, $2^n > 0$. Allora possiamo osservare che tale diseguaglianza è vera per $n = 0$. Supposta vera per n, supposto cioè che $2^n > 0$, risulta che $2 \cdot 2^n > 2 \cdot 0 = 0$ e quindi che la diseguaglianza è vera per $n+1$. Pertanto $2^n > 0$ per ogni $n \in N$.

Esempio. Supponiamo di voler dimostrare che

"la somma dei primi n numeri naturali è uguale a $n \cdot (n+1)/2$"

Per $n = 1$ l'asserzione è vera[6]. Supponiamo che l'asserzione sia vera per n. Allora, poiché

$$1 + \ldots + n + (n+1) = \frac{n(n+1)}{2} + (n+1)$$
$$= \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2},$$

l'asserzione è vera anche per $n+1$. Per il principio di induzione l'asserzione è vera per ogni n.

Esercizio. Trovare l'errore in questa dimostrazione.

"Teorema": Tutte le persone hanno la stessa età.

Dim. Indichiamo con $S(n)$ l'asserzione "in un gruppo di n persone tutte hanno la stessa età". Valgono le seguenti due asserzioni:

1: $S(1)$

2: $\forall n(S(n) \Rightarrow S(n+1))$.

Infatti la prima è ovvia, per provare la seconda fissiamo n e sia G un gruppo con $n+1$ persone. Affermo che due persone P_1 e P_2 in G hanno la stessa età. Infatti sia P una persona diversa da P_1 e P_2, poiché $G - \{P\}$ contiene n persone per ipotesi di induzione in $G - \{P\}$ tutte le persone hanno la stessa età. In particolare P_1 ha la stessa età di P_2.

Paradosso del mucchio di grano ed induzione:

Teorema: Se si accetta che non tutti i mucchi di grano sono piccoli allora il principio di induzione è falso.

[6] Il fatto che si può partire da 1 e non necessariamente da 0 sarà giustificato nel seguito.

Dim. Indichiamo con $P(n)$ l'asserzione "un mucchio con n chicchi è piccolo" allora valgono le seguenti asserzioni

1. $P(1)$ è vera
2. $\forall n(P(n) \to P(n+1))$ è vera.

Infatti la prima è indisctibile per la seconda basta osservare che se un mucchio di grano è piccolo, allora rimane piccolo anche se ci aggiungo un chicco di grano. D'altra parte se **2** fosse falsa esisterebbe $\underline{m}$ per cui $P(\underline{m}) \to P(\underline{m}+1)$ è falsa. Tenendo conto che una implicazione è falsa solo se l'antecedente è vero ed il conseguente è falso, avremmo che $P(\underline{m})$ sarebbe vera e $P(\underline{m}+1)$ falsa. In altre parole esisterebbe un mucchio magico il quale, pur essendo piccolo, diventa improvvisamente grande con l'aggiunta di un solo chicco.

Concludiamo osservando che se il principio di induzione fosse valido, da 1. e 2. seguirebbe la verità di $P(n)$ per ogni n in contrasto con le ipotesi. Quindi il principio di induzione è falso.

3. Definizione per ricorsione

Il principio di induzione permette di definire *"per ricorsione"* funzioni ed operazioni sui numeri naturali. Un esempio tipico è quello della la funzione fattoriale che indichiamo con *fatt.* Di solito tale funzione viene definita dicendo che, per ogni naturale n, $fatt(n)$ è il prodotto dei primi n numeri, oppure si scrive $fatt(n) = 1\cdot 2\cdot\ldots\cdot n$. Un modo più elegante e preciso di definire il fattoriale è dire che è l'unica funzione che soddisfa le condizioni

$$fatt(0) = 1 \quad ; \quad fatt(n+1) = fatt(n)\cdot(n+1).$$

Esercizio. Definire la funzione $f(n) = 2^n$ per ricorsione su n e dire perché la stessa funzione debba essere definita in modo diverso quando n varia nel campo dei razionali.

Per fare un altro esempio consideriamo il seguente problema.

Il problema delle strette di mano. Poniamoci il seguente problema:

Quante strette di mano devono darsi 3 persone che si incontrano in una festa ?

La risposta è semplice e diretta ed è il numero 3 (3 è abbastanza piccolo da permetterci di immaginare direttamente la scena delle strette di mano). Passiamo ora alla domanda:

Quante strette di mano devono darsi 7 persone che si incontrano in una festa ?

Si invita chi legge a dedicare un po' di tempo a risolvere questo problema. Si accorgerà che la risposta richiede un minimo di tempo e pazienza. La risposta diventa poi difficile se al posto di 7 si considera un numero più grande, ad esempio il numero 10.

Paradossalmente è più semplice invece affrontare il problema in generale e chiedersi:

Quante strette di mano devono darsi n persone che si incontrano in una festa ?

Se indichiamo con $f(n)$ tale numero possiamo tentare di calcolare i primi valori della funzione f. E' evidente che $f(1) = 0$. Infatti in un gruppo con una sola persona non possono esserci strette di mani. E' anche facile vedere che $f(2) = 1$ e che, come abbiamo già visto, $f(3) = 3$ ma già il calcolo di $f(4)$ si presenta un po' noioso . . .

Tuttavia un buon matematico dovrebbe accorgersi che l'avere calcolato $f(3)$ può essere utilizzato nel calcolo di $f(4)$. Infatti se inizialmente alla festa arrivano solo tre persone che si danno $f(3)$ strette di mano ed arriva una nuova persona, a tale nuova persona non resta che fare tre strette di mano. Quindi $f(4) = f(3)+3 = 6$. Se poi alle 4 persone si aggiunge un nuovo venuto, allora si devono aggiungere ancora quattro strette a quelle già fatte. Pertanto $f(5) = f(4)+4 = 10$. Più in generale, per rispondere alla nostra domanda basta calcolare i valori della successione

$f(1) = 0, f(2) = f(1)+1 = 1, f(3) = f(2)+2 = 3, f(4) = f(3)+3 = 6,$
$f(5) = f(4)+4 = 10, f(6) = f(5)+5 = 15, f(7) = f(6)+6 = 21,$
$f(8) = f(7)+7 = 28, f(9) = f(8)+8 = 36, f(10) = f(9)+9 = 45.$

Si osservi che la funzione f è completamente definita dalle due equazioni

$$f(1) = 0 \ ; \ f(n+1) = f(n)+n$$

che mostrano come, similmente a quanto avviene per il fattoriale, si possa calcolare il valore di f in un numero in funzione del valore di f in un numero precedente.[7] Questi due esempi suggeriscono la seguente definizione.

[7] E' facile vedere che $f(n) = 1+2+...+n$-1, cioè che $f(n)$ è la somma dei primi n-1 numeri naturali. Abbiamo già incontrato una funzione simile

Definizione 3.1. Sia (S,s,z_0) una terna di Peano ed $f : S \to S$ una funzione. Diciamo che f è definita per *ricorsione* tramite l'elemento $c \in S$ e la funzione $h : S^2 \to S$ se soddisfa le equazioni:

$$f(z_0) = c \ ; \ f(s(n)) = h(n, f(n)). \qquad (3.1)$$

La prima equazione in (3.1) viene detta *"assegnazione iniziale"* mentre la seconda *"schema di ricorsione"*. Lo schema di ricorsione dice che il valore di f in un numero può essere calcolato in funzione del valore di f nel precedente di tale numero. Nello schema di ricorsione compare la stranezza per cui si definisce una cosa utilizzando la cosa stessa, infatti si definisce f "ricorrendo" ad f stesso. Tuttavia si deve osservare che la funzione f a destra dello schema viene applicata ad un numero n che è più semplice del numero $s(n)$ che compare a sinistra. Pertanto applicando più volte lo schema si finisce con il dovere applicare f al valore z_0, cosa questa che viene consentita dall' assegnazione iniziale.

Fissati c ed h possiamo vedere le due equazioni in (3.1) come una "definizione" della funzione f anche se è necessario stare attenti all'uso dell'espressione "definizione". Infatti quando definiamo un ente matematico tramite una serie di proprietà allora la definizione è corretta solo se esiste un ed un solo ente verificante tale proprietà. Ad esempio se, nell'ambito della teoria dei numeri reali dico "sia r la radice di -2" ho una definizione non corretta poiché non esiste nessuna numero reale che soddisfa tale condizione. Se dico "sia r la radice di 2" ho una definizione non corretta poiché esistono due numeri reali che soddisfano tale condizioni. Una definizione corretta è invece, ad esempio, "sia r la radice positiva di 2". Ovviamente anche per la definizione 3.1 si pone la stessa questione. Da notare che in ambiti che non sono le terne di Peano tale teorema di esistenza e di unicità non è detto che valga. Ad esempio vale la seguente proposizione.

Proposizione 3.2. Nell'anello degli interi modulo m non esiste nessuna funzione *fatt* che verifica le due equazioni

$$fatt(0) = 1 \ ; \ fatt(x+1) = fatt(x)\cdot(x+1). \qquad (3.2)$$

In altre parole il fattoriale non si può definire.

---

quando abbiamo parlato dei numeri triangolari. Si prova, per induzione su n, che $f(n) = n\cdot(n-1)/2$.

Dim. Ad esempio consideriamo il caso di $m = 3$, e supponiamo per assurdo che esista una funzione *fatt* che verifica entrambe le condizioni (3.2). Proviamo allora a calcolare i valori di tale funzione:

$fatt(0) = 1$, (per la prima equazione)
$fatt(1) = fatt(0)\cdot 1 = 1$, (per la seconda equazione)
$fatt(2) = fatt(1)\cdot 2 = 2$, (per la seconda equazione).

Sembra quindi che non ci siano problemi. Tuttavia potremmo calcolare $fatt(0)$ anche servendoci della seconda equazione essendo 0 il successore di 2. In questo caso avremmo

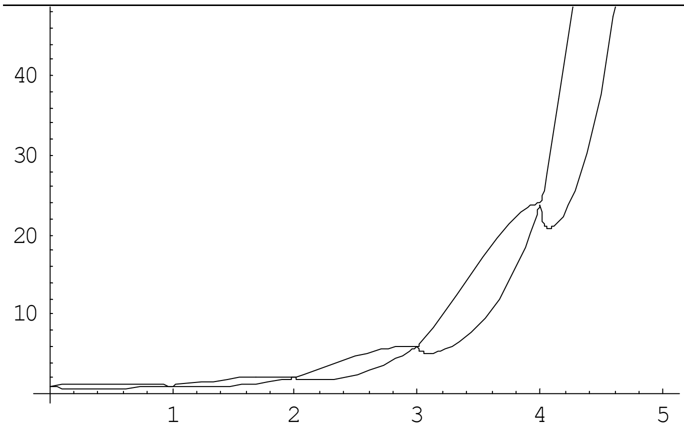
$fatt(0) = fatt(2)\cdot 0 = 0$

e quindi risulterebbe l'assurdo per cui la funzione *fatt* assume due valori diversi 0 ed 1.
La stessa cosa si può dire se m è un qualunque numero natural. Infatti, da un lato deve essere $fatt(0) = 1$, dall'altro deve essere $fatt(0) = fatt(m\text{-}1)\cdot 0 = 0$. □

Proposizione 3.3. Nell'insieme R^+ dei numeri reali maggiori o uguali a 0 esistono infinite funzioni che verificano le equazioni (3.2).

Dim. Sia g una qualunque funzione nell'intervallo [0,1] tale che $g(0) = 1$ e $g(1) = 1$. Allora posso definire $fatt : R^+ \to R^+$ in modo che la restrizione di *fatt* a [0,1] coincida con g e siano verificate le condizioni (3.2). Tale funzione è ovunque definita in quanto i valori definiti in un intervallo $[n,n+1]$ permettono di calcolare i valori nell'intervallo successivo. Ad esempio nell'intervallo [1,2] accade che poniamo $fatt(x) = fatt(x\text{-}1)\cdot x = g(x\text{-}1)\cdot x$, nell'intervallo [2,3] poniamo $fatt(x) = fatt(x\text{-}1)\cdot x = g(x\text{-}2)\cdot(x\text{-}1)\cdot x$ e così via. Poiché g può essere definito in infiniti modi diversi, esistono infinite funzioni in R^+ che verificano (3.2).

Nel seguente grafico consideriamo il caso in cui $g(x) = x\cdot(1\text{-}x)+1$ ed il caso in cui $g(x) = (x\text{-}1)\cdot\sqrt{x}+1$.



Grafici di due funzioni che verificano (3.2).

□

Nelle terne di Peano invece le cose funzionano bene. Vale infatti il seguente teorema.

Teorema 3.4. Sia (S, s, z_0) una terna di Peano, allora per ogni $c \in S$ ed $h : S^2 \to S$ esiste ed è unica una funzione f definita in tutto S che soddisfa le equazioni in (3.1). Tale funzione è totale.

Dim. Omettiamo la dimostrazione di esistenza che risulta essere un po' noiosa. Chi fosse interessato può vedere [Zanardo 2010].[8] Abbiamo visto tuttavia che tale esistenza non è verificata se ci si riferisce agli interi modulo m e che quindi le cose sono meno semplici di come appare.

Più facile provare l'unicità. Infatti supponiamo che f ed f' siano due funzioni soddisfacenti le equazioni (3.1) e sia $X = \{x \in S : f(x) = f'(x)\}$. Allora è evidente che $z_0 \in X$ e che se $n \in X$ allora, poichè

$$f(s(n)) = h(n, f(n)) = h(n, f'(n)) = f'(s(n)),$$

risulta che $s(n) \in X$. Pertanto $X = S$ e quindi $f = f'$. □

[8] D'altra parte nella maggior parte dei testi l'esistenza viene data come fatto ovvio (in generale nei testi di informatica). Infatti (3.1) rappresenta un algoritmo per effettuare un calcolo e tale algoritmo fornisce un unico output per ogni possibile input. Sembra naturale accettare che se esiste un algoritmo esiste anche la funzione corrispondente.

4. Esistenza ed "unicità" delle terne di Peano

Fino ad ora abbiamo parlato delle terne di Peano senza preoccuparci se tali terne esistano o meno. Infatti non basta fissare un sistema di assiomi per essere sicuri che esista un modello corrispondente di tale sistema. Per dimostrare tale esistenza è conveniente riformulare la nozione di terna di Peano in termini più algebrici.

Ricordiamo che viene chiamata *parte stabile* di una struttura algebrica A ogni sottoinsieme di A che contenga gli elementi designati e che sia chiuso rispetto alle operazioni della struttura. L'insieme delle parti stabili è un sistema di chiusura, cioè l'intersezione di una famiglia di parti stabili di una struttura è ancora una parte stabile. Ciò permette, dato un sottoinsieme X di A, di definire la *parte stabile generata da X* come l'intersezione di tutte le parti stabili contenenti X. Indichiamo con $<X>$ tale parte. Se $<X> = A$, allora si dice che X è un sistema di generatori di A. Ad esempio in un gruppo $(G, \cdot, ^{-1}, 1)$ una parte stabile è un sottoinsieme G' di G tale che

- G contiene 1
- il prodotto di due elementi di G' appartiene ancora a G',
- l'inverso di un elemento in G' appartiene ancora a G'.

In definitiva le parti stabili di G coincidono con i sottogruppi di G. Se G può essere generato da un solo elemento, allora si dice che G è ciclico.

Proposizione 4.1. Una struttura algebrica (S,s,z_0) soddisfacente $\boldsymbol{P_1}$ e $\boldsymbol{P_2}$ è una terna di Peano se e solo se ammette z_0 come generatore, cioè se S è la più piccola parte stabile contenente z_0.

Dim. Ovvia perché il principio di induzione afferma proprio che ogni parte X che sia stabile rispetto all'operazione s e che contenga z_0 coincide con S. □

Ricordiamo ora la definizione di insieme infinito.[9]

Definizione 4.2. Chiamiamo *infinito* un insieme T che sia equipotente ad una sua parte propria, cioè tale che esista una funzione iniettiva $f : T \to T$ che non sia suriettiva, cioè $f(T) \neq T$.

---

[9] Tale definizione, come quella di insieme finito, sarà esaminata meglio nel secondo volume e presenta più problemi di come possa apparire.

L'esistenza di una terna di Peano equivale ad accettare l'esistenza di un insieme infinito.

Teorema 4.3. Ammettere l'esistenza di una terna di Peano equivale ad ammettere l'esistenza di un insieme infinito.

Dim. Supponiamo che esista una terna di Peano (S,s,z_0), allora essendo la funzione successore iniettiva e poiché $z_0 \notin s(S)$, S è un insieme infinito. Viceversa sia T un insieme infinito e sia $f : T \to T$ una funzione iniettiva tale che $f(T) \neq T$. Allora esiste un elemento $z_0 \notin f(T)$ e possiamo prendere in considerazione la struttura algebrica (T, f, z_0). Sia $S = <z_0>$ la parte stabile generata da z_0, vogliamo provare che la sottostruttura (S,f,z_0) è una terna di Peano.[10] Infatti, gli assiomi $\boldsymbol{P_1}$ e $\boldsymbol{P_2}$ sono evidenti. Per provare il principio di induzione basta osservare che S, per costruzione, è la più piccola parte stabile contenente z_0. □

Avendo provato l'esistenza passiamo all'unicità delle terne di Peano. In realtà quando si ha una teoria T non ha senso parlare di unicità di modelli. Infatti non è difficile vedere che dato un modello di T esistono infinite copie isomorfe di tale modello e che tali copie sono ancora modelli di T. Allora si è costretti a ricorrere ad una nozione più debole di quella di unicità che è quella di *"unicità a meno dei isomorfismi"*.

Definizione 4.4. Una teoria T si dice *categorica* se tutti i modelli di tale teoria sono isomorfi tra loro.

Se una teoria è categorica si dice anche che *esiste un unico modello a meno di isomorfismi* di tale teoria. Ad esempio, la teoria dei gruppi non è categorica poiché, come è noto, esistono gruppi che non sono isomorfi tra loro. La teoria dei gruppi di ordine 7 è invece categorica in quanto tutti i gruppi di ordine 7 sono isomorfi tra loro.

Ricordiamo che un omomorfismo tra due strutture algebriche è una funzione che porta elementi designati in elementi designati e "conserva" tutte le operazioni (l'immagine del composto è il

[10] Per la nozione di sottostruttura generata si veda l'Appendice.

composto delle immagini). Nel caso di due terne di Peano (S,s,z_0) e (S',s',z_0') un omomorfismo è una funzione $f : S \rightarrow S'$ tale che,

$$f(z_0) = z_0' \quad ; \quad f(s(x)) = s(f(x)).$$

Teorema 4.5. La teoria delle terne di Peano è categorica. Precisamente, date due terne di Peano esiste uno ed un solo isomorfismo tra le due terne.

Dim. Siano (S,s,z_0) e (S',s',z_0') due terne di Peano. Allora per trovare un omomorfismo dobbiamo trovare una funzione $f : S \rightarrow S'$ tale che,

$$f(z_0) = z_0' \quad ; \quad f(s(x)) = s(f(x)). \tag{2}$$

Ma tali condizioni costituiscono una definizione per ricorsione e quindi, per quanto abbiamo visto nel Teorema 3.4, esiste una ed una sola funzione f che soddisfa tali condizioni. Tale funzione per il modo in cui è stata definita è un omomorfismo. Si prova facilmente che f è un isomorfismo.[11] □

Poiché tutte le terne di Peano sono isomorfe tra loro, non ha importanza quale di esse viene considerata. D'ora in poi supponiamo che ne sia stata fissata una, indicheremo con N_0 il suo dominio e con 0 il suo primo elemento e chiamiamo *numero naturale* ogni elemento di N_0. Con N si denota l'insieme N_0-$\{0\}$, cioè l'insieme dei numeri naturali diversi da zero.

5. Addizione, moltiplicazione, ordine

E' possibile estendere la nozione di definizione per ricorsione in modo da poter definire funzioni di più variabili. Ad esempio, consideriamo la funzione potenza n-esima di base b con n numero naturale e b numero reale, consideriamo cioè la funzione $pot(b,n) = b^n$. In questo caso $pot : R \times N_0 \rightarrow R$ è l'unica funzione che soddisfa le condizioni

$$pot(b, 0) = 1 \quad ; \quad pot(b,n+1) = pot(b,n) \cdot b.$$

Tale esempio suggerisce la seguente, più generale, definizione.

[11] Nel caso in cui le due strutture (S,s,z_0) ed (S',s',z_0') coincidono questo teorema ci dice che esiste uno ed un solo automorfismo in una terna di Peano. Questo automorfismo non può che essere la funzione identica. Quando il gruppo degli automorfismi coincide con il sottogruppo costituito dalla sola applicazione identica, allora si dice anche che la struttura *rigida*. Le terne di Peano sono esempi di strutture rigide.

Definizione 5.1. Sia (S,s,z_0) una terna di Peano ed A e B insiemi non vuoti. Allora diciamo che la funzione n-aria $f : A\times S \rightarrow B$ è definita per *ricorsione sulla seconda variabile* se esistono due funzioni $g : A{:}\rightarrow B$ ed $h : A\times S^2 \rightarrow B$ tali che:

$$f(a,z_0) = g(a) \;\; ; \;\; f(a, s(n)) = h(a, n, f(a,n)). \qquad (5.1)$$

Da notare che in tale definizione si richiede che solo uno degli insiemi sia una terna di Peano. Con tale nozione estesa di ricorsione è possibile definire in una terna di Peano le operazioni base di addizione e moltiplicazione. Infatti la somma è definita come applicazione iterata del successore, il prodotto come applicazione iterata della somma.

Definizione 5.2. In una terna di Peano $(N_0, s, 0)$ chiamiamo *addizione* la funzione *som* : $N_0\times N_0\rightarrow N_0$ definita per ricorsione tramite le due equazioni:

$$som(x,0) = x \quad ; \quad som(x,s(y)) = s(som(x,y)). \qquad (3)$$

Chiamiamo *moltiplicazione* la funzione *pro* : $N_0\times N_0\rightarrow N_0$ definita per ricorsione dalle due equazioni

$$pro(x,0) = 0 \quad ; \quad pro(x,s(y)) = som(pro(x,y),x). \qquad (4)$$

In generale la funzione addizione viene indicata con il simbolo + e la funzione di moltiplicazione con un puntino $\cdot$. Inoltre si preferiscono le notazioni *infisse* $x+y$ e $x\cdot y$ al posto delle notazioni *prefisse* $som(x,y)$ e $pro(x,y)$.[12] Se denotiamo con 1 l'elemento $s(0)$, allora risulta che $som(x,1) = som(x,s(0)) = s(som(x,0)) = s(x)$. Pertanto, utilizzando la notazione additiva, possiamo indicare con $x+1$ il successore di x. Per le operazioni ora definite valgono le seguenti proprietà di cui omettiamo la dimostrazione.

[12] Da notare che se provassimo ad estendere ai numeri reali le definizioni ora date di addizione e di moltiplicazione apparirebbero subito delle difficoltà in quanto il campo dei numeri reali non è una terna di Peano. Le difficoltà riguarderebbero sia il processo di calcolo, sia l'unicità della funzione definita. Ad esempio il tentativo di calcolare $som(1,2.5)$ condurrebbe a calcolare $som(1, 1.5)$, e quindi $som(1, 0.5)$ e quindi $som(1, -0.5)$ e poi $som(1, -1.5)$ e così all'infinito. Per quanto riguarda l'unicità la situazione è la stessa di quella già osservata per il fattoriale.

Proposizione 5.3. Le operazioni + e · sono associative e commutative ed ammettono come elemento neutro 0 ed 1, rispettivamente. Inoltre vale la proprietà distributiva del prodotto rispetto la somma.

Poniamoci ora il problema di come definire una relazione d'ordine in una terna di Peano. Naturalmente le possibili relazioni d'ordine sono infinite, ad esempio potrei considerare la relazione $\lesssim$ tale che $n \lesssim m$ se m è un multiplo di n. Tuttavia è ragionevole trovare una definizione che rispetti la nostra intuizione secondo cui

- il successivo di un numero sia maggiore del numero stesso, cioè risulti che $x \leq s(x)$.

D'altra parte se questo avviene per ogni x deve essere anche $s(x) \leq s(s(x))$ e quindi, per la proprietà transitiva delle relazioni d'ordine, $x \leq s^2(x)$. Più in generale per la relazione che stiamo cercando deve valere la condizione $x \leq s^n(x)$ dove, per ogni $n \in N_0$ la funzione s^n è definita, per ricorsione su n, tramite le equazioni

$$s^0(x) = x \ ; \ s^{n+1}(x) = s(s^n(x)).$$

Ciò suggerisce la seguente definizione.

Definizione 5.4. Definiamo la relazione $\leq$ ponendo

$$x \leq y \Leftrightarrow \text{esiste } n \in N_0 \text{ tale che } s^n(x) = y.$$

In altre parole diciamo che x è minore o uguale ad y o nel caso in cui $x = y$ oppure se partendo da x dopo un numero finito di passi si raggiunge y. Per potere dimostrare che $\leq$ è una relazione d'ordine abbiamo bisogno del seguente lemma.

Lemma 5.5. Fissato $n \neq 0$, per ogni $x \in N_0$ risulta che

$$x \neq s^n(x). \tag{5.2}$$

Dim. Si procede per induzione su x. Infatti (5.2) risulta vera per $x = 0$ in quanto 0 non è successore di nessun elemento. Supponiamo che (5.2) sia vera per x, cioè che $x \neq s^n(x)$. Allora, poiché s è una funzione iniettiva, sarà anche $s(x) \neq s(s^n(x)) = s^n(s(x))$. Pertanto (5.2) è vera anche per $s(x)$. Per il principio di induzione possiamo concludere che (5.2) è vera per ogni $x \in N_0$. □

Possiamo ora provare il seguente teorema che giustifica la scelta fatta nel definire la relazione d'ordine in una terna di Peano.

Teorema 5.6. La relazione $\leq$ è la più piccola relazione d'ordine tale che, per ogni $x \in N_0$, $x \leq s(x)$.

Dim. Poiché $s^0(x) = x$ abbiamo che $x \leq x$ e quindi $\leq$ verifica la proprietà riflessiva. Inoltre

$$x \leq y,\ y \leq z \ \Rightarrow \exists n,m\ \ y = s^n(x) \text{ e } z = s^m(y) \Rightarrow z = s^{n+m}(x) \Rightarrow x \leq z.$$

e questo prova che $\leq$ è una relazione transitiva. Per provare che vale la proprietà anti-simmetrica, osserviamo che

$$x \leq y,\ y \leq x \ \Rightarrow \exists n,m\ \ y = s^n(x) \text{ e } x = s^m(y)$$
$$\Rightarrow x = s^m(s^n(x)) \Rightarrow x = s^{m+n}(x) \Rightarrow n = m = 0 \ \Rightarrow\ x = y.$$

Abbiamo pertanto provato che $\leq$ è una relazione d'ordine. Per il modo in cui abbiamo definito tale relazione, è immediato che essa contiene l'insieme $\{(x,s(x)) : x \in N_0\}$ di coppie, cioè che $x \leq s(x)$. Per provare che $\leq$ è la più piccola relazione d'ordine che soddisfa tale proprietà, supponiamo che $\mathcal{R}$ sia una relazione di ordine contenente $\{(x,s(x)) : x \in N_0\}$. Vogliamo provare che $\mathcal{R}$ contiene $\leq$, cioè che, fissato x,

$$(x, s^n(x)) \in \mathcal{R} \text{ per ogni } n. \tag{5.3}$$

Procediamo per induzione su n. E' immediato che (5.3) è vera per $n = 0$. Supponiamo che (5.3) sia verificata da n, allora poiché $(s^n(x),s(s^n(x))) \in \mathcal{R}$, per la proprietà transitiva di $\mathcal{R}$ possiamo anche affermare che $(x, s^{n+1}(x)) \in \mathcal{R}$ e quindi (5.3) è vera per $n+1$. Ciò prova che $\mathcal{R}$ contiene $\leq$. □

Tale teorema ci permette di chiamare *relazione d'ordine naturale* la relazione $\leq$ e di dire che essa è la relazione d'ordine generata dalla funzione successore.

Studiamo ora alcune proprietà della relazione $\leq$.

Lemma 5.7. La relazione d'ordine naturale è una relazione totale il cui minimo è 0. Inoltre, dato $x \in N_0$, non esiste nessun elemento tra x e $s(x)$.

Dim. Per provare che $0 \leq x$ per ogni x, osserviamo che tale diseguaglianza vale per $x = 0$ e che se vale per x vale ovviamente anche per $s(x)$. Per il principio di induzione essa vale per ogni x.

Per provare che $\leq$ è totale consideriamo, dato un elemento x, l'insieme $Conf(x) = \{y \in N_0 : x \leq y \text{ oppure } y \leq x\}$ degli elementi

confrontabili con x. Vogliamo provare per induzione che $Conf(x) = N_0$. E' evidente che $0 \in Conf(x)$. Supponiamo che $y \in Conf(x)$, allora:
- caso $x \leq y$ per definizione di $\leq$ esiste $n \in N_0$ tale che $y = s^n(x)$ e quindi $s(y) = s^{n+1}(x)$. Ma ciò significa che $s(y) \in Conf(x)$.
- caso $y < x$, per definizione esiste $n \in N_0$ tale che $x = s^n(y)$ con $n \neq 0$, allora $s^{n-1}(s(y)) = x$ e quindi $s(y) \leq x$. Pertanto $s(y) \in Conf(x)$.

Per il principio di induzione $Conf(x) = N_0$ e quindi ogni elemento di N_0 è confrontabile con x. Infine, supponiamo per assurdo che esista x' tale che $x < x' < s(x)$ e quindi $n \neq 0$ ed $m \neq 0$ tali che $x' = s^n(x)$ e $s(x) = s^m(x')$. Allora $s(x) = s^m(s^n(x))$ e quindi $x = s^{m+n-1}(x)$ in contrasto con il lemma 5.5. □

Come prova il seguente teorema, $\leq$ è un buon ordine.[13]

Teorema 5.8. La relazione d'ordine naturale definita in una terna di Peano è una relazione di buon ordine.

Dim. Sia X un sottoinsieme non vuoto di N_0. Per provare che ammette un minimo indichiamo con $mino(X)$ l'insieme dei minoranti di X. Ovviamente $0 \in mino(X)$ e quindi se ogni elemento $m \in mino(X)$ avesse il successivo in $mino(X)$, per il principio di induzione avremmo che $mino(X) = N_0$, cioè tutti gli elementi sono minoranti di X. In questo caso, detto $\underline{m}$ un elemento di X avremmo che tutti gli elementi di N_0 sono minori di $\underline{m}$. Ciò è assurdo in quanto il successivo di $\underline{m}$ non può essere minore di $\underline{m}$. Questo prova che esiste un elemento $m \in mino(X)$ tale che $s(m) \notin mino(X)$. Affermo che m è il minimo di X. Infatti poiché $s(m)$ non è un minorante di X, esiste $x' \in X$ tale che $s(m)$ non è minore di x'. Poiché $\leq$ è un ordine totale, $x' < s(m)$ e quindi anche $m \leq x' < s(m)$. Per quanto detto nel lemma 5.7, ciò implica che $m = x' \in X$ e quindi che m è il minimo di X. □

Possiamo ora dare la seguente fondamentale definizione.

Definizione 5.9. Chiamiamo *sistema di numeri naturali* la struttura algebrica ordinata $(N_0, +, \cdot, 0, \leq)$ che si ottiene definendo in

[13] Una relazione viene detta di "buon ordine" se è una relazione d'ordine tale che ogni sottoinsieme non vuoto ammette il minimo. Le relazioni di buon ordine saranno studiate nel secondo volume.

una terna di Peano (N_0, s, 0) le operazioni di addizione e moltiplicazione e la relazione d'ordine ≤.[14]

6. Variazioni sul principio di induzione

La definizione della relazione d'ordine permette di riformulare in vari modi il principio di induzione. Ad esempio non è necessario "partire da zero" ma è possibile partire da qualunque intero.

Proposizione 6.1. (Principio di induzione a partire da un intero qualsiasi)**.** Sia P una proprietà definita in una terna di Peano (S, s, z_0) e $\underline{m} \in S$. Allora

se *a*) $P(\underline{m})$

e *b*) $\forall x \geq \underline{m}\ (P(x) \Rightarrow P(s(x)))$

allora *c*) $\forall x \geq \underline{m}\ P(x)$.

Dim.[15] Consideriamo la proprietà

$$P'(x) \equiv \text{"}x \leq \underline{m} \text{ oppure } P(x)\text{"}.$$

Allora $P'(x)$ verifica le due condizioni nella terna di Peano che permettono di affermare che P' vale per ogni x in S. Infatti P' è vera per z_0 (in quanto $z_0 \leq \underline{m}$). Supponiamo P' vera per x, allora dobbiamo provare che P' è verificata da $s(x)$. Ora se $x \geq \underline{m}$ per *b*) $s(x)$ verificando P verifica anche P'. Se invece $x < \underline{m}$ allora $s(x) \leq \underline{m}$ e quindi è evidente che anche $s(x)$ verifica P'. In conclusione per il principio di induzione risulta che per ogni x vale $P'(x)$ e quindi che vale *c*).

Equivalentemente possiamo porre

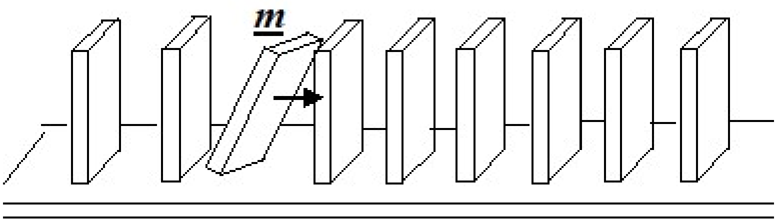
[14] Un modo totalmente diverso di introdurre la struttura algebrica dei numeri naturali è quello che si basa sulla nozione di numero cardinale finito. Questo punto di vista, che si basa sulla nozione di equipotenza, verrà esposto nel secondo volume.

[15] Una dimostrazione diversa di tale proposizione è che la funzione $f(x) = x + \underline{m}$ è un isomorfismo tra la terna di Peano (S,s,z_0) e la struttura $(\{x : x \geq \underline{m}\}, s, \underline{m})$. Poiché strutture isomorfe verificano le stesse proprietà, questo significa che questa seconda struttura è a sua volta una terna di Peano e quindi che per essa vale il principio di induzione. Naturalmente per rendere rigorosa tale dimostrazione si dovrebbe definire meglio cosa si intenda per "proprietà" e dimostrare che strutture isomorfe verificano le stesse proprietà. Tuttavia anche se apparentemente più complicato, tale modo di procedere permette di avere una visione più "dall'alto" e più profonda della matematica.

$$X = \{n\in S : n \text{ verifica } P\} \text{ e } D = X\cup\{x\in S : x<\underline{m}\}$$

Allora D contiene z_0 ed è tale che $s(D) \subseteq D$. Pertanto $D = S$ e quindi $X \supseteq \{x\in S : x\geq \underline{m}\}$. □

Questa forma del principio di induzione può essere visualizzata dalla seguente figura il cui significato è ovvio:



Problema. Supponiamo che in un tavolo di poker i giocatori abbiano ha a disposizione solo gettoni che valgono 3 oppure 5 euro (in quantità non limitata). Supponendo che il piatto sia inizialmente vuoto, dimostrare per induzione che il giocatore può fare qualunque tipo di puntata.

Un'altra variazione del principio di induzione che non coinvolge il successore ma solo l'ordine è la seguente.

Proposizione 6.2. (Principio di induzione transfinita). Sia P una proprietà definita in una terna di Peano, allora

se
$$(\forall x<n\ P(x)) \Rightarrow P(n) \quad (a^*)$$
allora $\forall n P(n)$.

Dim. Supponiamo che valga a^*) e che, per assurdo, esista un intero n per cui $P(n)$ sia falsa. Allora esiste il minimo $\underline{n}$ tra tali interi per cui succede questo. Ora, da un lato $P(\underline{n})$ risulta falsa, da un altro lato, essendo $\forall x<\underline{n}\ P(x)$, per a^*) $P(\underline{n})$ è vera. L'assurdo a cui siamo pervenuti ci assicura che $P(n)$ è vera per ogni n. □

Tale dimostrazione utilizza solo l'ipotesi che la relazione $\leq$ sia di buon ordine. Quindi può essere estesa a tutti gli insiemi con un buon ordine e qualunque sia la cardinalità di tali insiemi. Questo è il motivo per cui prende il nome di principio di induzione transfinita. Riconsidereremo tale principio nel secondo volume.

Il seguente teorema è un esempio di utilizzazione del principio di induzione transfinita.

Teorema 6.3. (Teorema fondamentale dell'aritmetica) Ogni numero naturale n diverso da zero può essere scomposto in un unico modo nel prodotto di un numero finito di primi.

Dim. Proviamo che n si può scrivere come prodotto di un numero finito di primi. Potremmo tentare di utilizzare il principio di induzione, ma in tale caso dovremmo provare che se n è scomponibile allora $n+1$ è ancora scomponibile. Ciò non sembra facile perché si dovrebbe ricavare da una scomposizione di n una scomposizione di $n+1$. Proviamo invece ad utilizzare il principio di induzione transfinita e supponiamo pertanto che tutti gli interi strettamente minori di n ammettono una scomposizione. Per provare che n ammette una scomposizione osserviamo che se n è primo allora banalmente ammette una scomposizione in fattori primi. Se n non è primo sarà uguale al prodotto di due interi a e b strettamente minori di n i quali pertanto, per ipotesi di induzione, sono scomponibili. Pertanto n, essendo prodotto di due numeri scomponibili, è ancora scomponibile. □

Ecco una ulteriore riformulazione del principio di induzione.

Proposizione 6.4. (Principio di induzione sul decorso dei valori). Sia P una proprietà definita in una terna di Peano, allora

se

$P(0)$ *a*)

e

$\forall\, x \leq n P(x) \Rightarrow P(n+1))$ *b*)

allora $\forall n P(n)$.

Dim. E' sufficiente provare che da *a*) e *b*) segue *a**). Infatti da *b*) segue che a^*) vale per ogni naturale del tipo $n+1$, cioè per ogni naturale successore. Per mostrare che vale anche per 0, cioè che $(\forall x<0\ P(x)) \Rightarrow P(0)$ basta osservare che per *a*) il conseguente di tale implicazione è vero. □

Tale principio viene indicato anche con altri nomi, ad esempio *"Seconda forma del principio di induzione"* oppure *"Principio di induzione completo"*.

7. Ogni struttura algebrica ha un suo principio di induzione
Se, come abbiamo fatto nel paragrafo 4, si guarda ad una terna di Peano in termini algebrici, si vede subito che il principio di induzione rientra in un principio più generale.

Teorema 7.1. Sia $(D, h_1,..., h_t, z_0,..., z_m)$ una struttura algebrica avente $z_0,...,z_m$ come sistema di generatori. Supponiamo inoltre che P sia una proprietà tale che:
a) P è verificata da $z_0,...,z_m$
b) per ogni operazione n-aria h_i se P è verificata da $x_1,...,x_n$ allora è verificata da $h_i(x_1,...,x_n)$
allora P è verificata per ogni $x \in D$.

Dim. Evidente. Infatti l'insieme degli elementi di D che verificano P è la più piccola parte stabile che contiene gli elementi $z_0,\dots,z_m$. D'altra parte l'ipotesi che $z_0,...,z_m$ sono un sistema di generatori comporta che tale parte stabile coincida con D. □

Esempio. Proponiamoci di dimostrare che un numero naturale è divisibile per 2 se e solo se l'ultima cifra nella sua rappresentazione decimale è tra 0, 2, 4, 6, 8.

Sia P_0 l'insieme dei numeri pari compreso 0 e definiamo l'operazione p ponendo $p(n) = s(s(n))$. Allora la struttura algebrica $(P_0, p, 0)$ ha come generatore 0. Indichiamo con $D(n)$ la proprietà "n termina con una delle cifre 0, 2, 4, 6, 8". Allora è semplice verificare che

- $D(0)$
- $D(x) \Rightarrow D(p(x))$.

Pertanto D è verificata in ogni elemento di P_0. Viceversa sia n un numero naturale che verifica D, allora poiché n si rappresenta come somma di potenze di dieci (tutte divisibili per 2) più la parte intera che abbiamo supposto pari, n è pari.

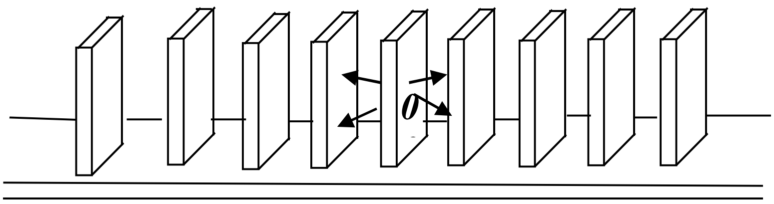
In modo analogo possiamo procedere per provare che un numero è divisibile per 5 se e solo se termina per 0 o per 5.

Esempio: Il *"principio di induzione" per Z.* Come già osservato, la struttura $(Z,s,0)$, dove $s(x) = x+1$, non è una terna di Peano e quindi in Z non sarebbe possibile fare dimostrazioni per induzione. Tuttavia se vogliamo provare che una proposizione vale per ogni elemento di Z possiamo riferirci alla struttura algebrica $(Z, s_+, s_-, 0)$ dove $s_+(x) = x+1$ e $s_-(x) = x-1$. Infatti la sottostruttura di $(Z, s_+, s_-, 0)$ generata da 0 coincide con $(Z, s_+, s_-, 0)$, cioè 0 è un generatore di $(Z, s_+, s_-, 0)$. Se volessimo formulare un principio di induzione per Z dovremmo quindi dire che, per ogni proprietà P definita in Z,

se $P(0)$
e $P(x) \Rightarrow P(x+1)$ e $P(x-1)$,
allora P è verificata per ogni $x \in Z$.

In tale caso l'applicazione del principio di induzione può essere visualizzata accettando che i pezzi del gioco domino siano poggiati su di un tavolo infinito sia a sinistra che a destra. Supponiamo che tali pezzi siano esplosivi e che:
- il pezzo **0** esploda
- se un pezzo x esplode allora fa esplodere i due pezzi vicini, cioè sia il pezzo successivo $s_+(x)$ che quello precedente $s_-(x)$,

allora è evidente che tutti i pezzi esplodono.



Esempio: Il *"principio di induzione"* in Z/m. Abbiamo già osservato che la struttura $(Z/m, s, [0])$, dove s è definita ponendo $s([n]) = [n+1]$ non è una terna di Peano. Allo stesso tempo è vero che [0] è un generatore di tale struttura poiché ogni intero modulo m si può ottenere a partire da [0] applicando un certo numero di volte s. Pertanto pur non essendo $(Z/m, s, [0])$ una terna di Peano, per tale struttura continua a valere il principio di induzione.

8. L'anello degli interi relativi

I numeri naturali sono uno strumento per misurare la grandezza di insiemi finiti. Ad esempio la terna di Peano che corrisponde alle possibili tacche su di un pezzo di legno può avere come scopo il contare il numero di pecore o il numero dei giorni passati od altro. Tuttavia esistono tipi di attività in cui i numeri naturali si mostrano non essere adeguati. Supponiamo ad esempio di dovere distinguere in una contabilità i soldi che devono essere dati dai soldi che si devono ricevere da alcuni clienti. Se la contabilità è tenuta su due colonne avremo una situazione del tipo indicato nella seguente tabella.

Cliente	Aver e	Dar e
Carlo	7	5
Luigi	3	6
Maria	10	8
. . .	. . .	. . .

Ne segue che l'informazione relativa a Carlo è rappresentata dalla coppia (7,5), l'informazione relativa a Luigi è rappresentata dalla coppia (3,6), quella relativa a Maria da (10,8). Ciò suggerisce l'introduzione di un nuovo tipo di numero costituito da due parti (quindi una coppia) ciascuna con un proprio significato. Naturalmente oltre all'interpretazione di una coppia in termini di debiti e crediti, sono possibili diverse altre interpretazioni. Ad esempio possiamo interpretare una coppia (n,m)
- come l'operazione "aggiungere n e togliere m".
- come "fare n passi avanti ed m passi indietro"
- come "applicare una forza di grandezza n in una direzione ed una forza di grandezza m nella direzione opposta".

In definitiva partiamo dall'insieme dei numeri naturali N_0 e consideriamo l'insieme il prodotto cartesiano $N_0\times N_0$. In tale insieme definiamo opportune operazioni.

Proposizione 8.1. Sia $(N_0\times N_0, +, \cdot, \mathbf{0}, \mathbf{1})$ la struttura algebrica in cui l'addizione + è definita ponendo:

$$(n,m) + (a,b) = (n+a, m+b), \tag{8.1}$$

e l'addizione $\cdot$ è definita ponendo:

$$(n,m) \cdot (a,b) = (na+mb, nb+ma). \tag{8.2}$$

Inoltre **0** = (0,0), **1** = (1,0).

Il motivo per cui la somma viene definita in questo modo è ovvio. Se si fanno n passi avanti ed m indietro e poi si fanno a passi avanti e b indietro, allora globalmente si sono fatti $n+a$ passi avanti ed $m+b$ indietro. Giustificare la definizione di prodotto non è altrettanto facile perché, ad esempio, non ha molto senso moltiplicare passi avanti con passi indietro (un problema analogo si presenta quando si deve definire la moltiplicazione tra numeri complessi). Probabilmente non esiste una giustificazione della moltiplicazione che sia "semantica" cioè che si riferisca ad operazioni concrete. Se si vuole trovare la reale giustificazione si deve guardare al livello sintattico, cioè all'esigenza di costruire un ambiente "comodo per i calcoli". Questo significa che la moltiplicazione deve essere associativa, commutativa, deve valere la proprietà distributiva e così via.

Proposizione 8.2. Data la struttura $(N_0\times N_0, +, \cdot, 0, 1)$ le operazioni + e · sono associative e commutative ed ammettono 0 = (0,0) e 1 = (1,0) come elemento neutro, rispettivamente. Inoltre vale la proprietà distributiva. Tuttavia $(N_0\times N_0, +, 0)$ non è un gruppo e quindi $(N_0\times N_0, +, \cdot, 0, 1)$ non è un anello.

Dim. Ci limitiamo a mostrare che $(N_0\times N_0, +, 0)$ non è un gruppo. Infatti, dato un elemento (m,n) diverso da (0,0), qualunque sia (a,b) risulta che

$$(m,n) + (a,b) = (m+a, n+b) \neq (0,0).$$

Pertanto in $(N_0\times N_0, +, \mathbf{0})$ non esiste l'opposto di (m,n). □

La struttura $(N_0\times N_0, +, \cdot, \mathbf{0}, \mathbf{1})$ non solo non è un anello ma non risulta nemmeno adeguata a rappresentare le situazioni da cui siamo partiti all'inizio del paragrafo. Infatti elementi diversi di $N_0\times N_0$ possono rappresentare la stessa situazione. Ad esempio, se si interpretano le coppie in termini di debiti e di crediti, allora è naturale considerare (n,m) equivalente a (n',m') se avere n e dare m risulta equivalente ad avere n' e dare m'. Ad esempio se si guarda la tabella, è evidente che Carlo e Maria sono due clienti con la stessa situazione finanziaria, cioè che la coppia (7,5) è equivalente alla coppia (10,8). Ancora, è naturale considerare (n,m) equivalente a (n',m') se fare n passi avanti ed m indietro produce lo stesso risultato di fare n' passi avanti ed m' indietro.

Potremmo allora dire che (n,m) è equivalente a (n',m') se nel caso $n \geq m$ risulta che $n' \geq m'$ e $n - m = n'-m'$, mentre nel caso $n<m$ risulta che $n'<m'$ e $m-n = m'-n'$.

Proposizione 8.3. Definiamo la relazione $\equiv$ in $N_0 \times N_0$ ponendo,

$$(n,m) \equiv (n',m') \Leftrightarrow n+m' = m+n'. \tag{8.3}$$

Allora tale relazione è una congruenza della struttura algebrica $(N_0 \times N_0, +, \cdot, (0,0), (1,0))$.

Dim. Proviamo la compatibilità rispetto alla somma, cioè che

$$(n,m) \equiv (n',m')\ ,\ (a,b) \equiv (a',b') \Rightarrow (n+a,m+b) \equiv (n'+a',m'+b').$$

Infatti per ipotesi $n+m' = m+n'$ e $a+b' = b+a'$, da cui, sommando termine a termine, $n+m'+a+b' = m+n'+b+a'$ che equivale a $(n+a,m+b) \equiv (n'+a',m'+b')$.

La dimostrazione della compatibilità rispetto al prodotto si effettua in maniera analoga. □

Proposizione 8.4. La struttura quoziente di $(N_0 \times N_0, +, \cdot, 0, 1)$ modulo $\equiv$ è un anello unitario che chiamiamo *anello degli interi relativi* che estende la struttura $(N_0, +, \cdot, 0, 1)$. Indichiamo con $\boldsymbol{Z} = (Z, +, \cdot, 0, 1)$ tale anello.

Dim. Ci limitiamo ad osservare che ogni elemento $[(n,m)]$ ammette come opposto $[(m,n)]$. Infatti

$$[(n,m)]+[(m,n)] = [(n+m,m+n)] = [(0,0)].$$

Inoltre la funzione $f: N_0 \to Z$ definita ponendo $f(n) = [(n,0)]$ è una immersione di $(N_0, +, \cdot, 0, 1)$ in $(Z, +, \cdot, 0, 1)$. □

Pertanto $\boldsymbol{Z}$ è definita dalle equazioni

$$[(n,m)] = \{(n',m') \mid (n',m') \equiv (n,m)\}$$

$$[(n,m)]+[(a,b)] = [(n+a,m+b)]$$
$$[(n,m)] \cdot [(a,b)] = [(na+mb, nb+ma)]$$
$$0 = [\mathbf{0}],\ \ 1 = [\mathbf{1}].$$

Problema: Tenendo conto del modo come abbiamo definito $\boldsymbol{Z}$ dire che cosa significa che $-(-5) = 5$. $Z = \{[(n,m)] \mid (n,m) \in N_0 \times N_0\}$

Infine la seguente proposizione, di cui omettiamo la dimostrazione, mostra come si possa definire una relazione d'ordine in Z.

Proposizione 8.5. Definiamo in $N_0 \times N_0$, la relazione $\leq$ ponendo $(n,m) \leq (p,q)$ se e solo se $n+q \leq m+p$. Allora tale relazione è compatibile con $\equiv$ e definisce sul quoziente Z una relazione d'ordine $\leq$. La struttura $(Z, +, \cdot, 0,1, \leq)$ è un anello ordinato.

Esercizio. Provare che $[(1,0)]$ è l'elemento neutro rispetto al prodotto in Z.
Esercizio. Dimostrare che $(Z, +, 0)$ è un gruppo.
Esercizio. Dire perché è sbagliato definire l'operazione $\otimes$ ponendo $[(n,m)] \otimes [(a,b)] = [(na, mb)]$.

Concludiamo osservando che mentre i numeri naturali sono stati definiti col metodo assiomatico, i numeri interi relativi sono stati "costruiti" a partire dai numeri naturali. La stessa cosa avverrà per i razionali e per i reali.

9. Il campo dei razionali.
In modo analogo possiamo ottenere il passaggio dall'anello Z degli interi relativi al campo Q dei numeri razionali. In questo caso consideriamo l'insieme

$$Z \times (Z - \{0\}) = \{(p,q) \mid p \in Z, q \in Z, q \neq 0\}$$

e l'interpretazione che ora diamo ad una coppia (p,q) è "moltiplicare per p e dividere per q" ed usualmente è indicata con p/q. Chiamiamo *frazioni* le espressioni del tipo p/q. In tale insieme di coppie introduciamo due operazioni tramite le eguaglianze

$$(p,q)+(a,b) = (pb+qa,qb) \; ; \; (p,q)\cdot(a,b) = (pa,qb). \qquad (9.1)$$

In termini di frazioni poniamo

$$p/q+a/b = (pb+qa)/(qb) \; ; \; (p/q)\cdot(a/b) = (pa)/(qb).$$

In tale modo viene definita una struttura algebrica in $Z \times (Z - \{0\})$.

Proposizione 9.1. Nella struttura $(Z \times (Z-\{0\}), +, \cdot, (0,1), (1,1))$ le operazioni sono commutative ed associative, $(0,1)$ è elemento neutro rispetto all'addizione, $(1,1)$ è elemento neutro rispetto alla moltiplicazione. Tuttavia tale struttura non è un campo.

Dim. Poichè $(p,q)+(0,1) = (p1+q0, q1) = (p,q)$, la coppia $(0,1)$ è elemento neutro rispetto la somma. In modo simile si provano le altre proprietà. Per provare che la struttura non è un campo osserviamo che se una coppia (p,q) ammettesse inverso allora esisterebbero due interi x ed y in Z tali che $(p,q)\cdot(x,y) = (1,1)$, si avrebbe pertanto che $px = 1$ e $qy = 1$, e quindi p e q sarebbero invertibili in Z. Poiché gli unici elementi invertibili di Z sono 1 ed il suo opposto -1, ne segue che gli unici elementi invertibili della struttura $(Z\times(Z-\{0\}), +, \cdot,(0,1),(1,1))$ sono le coppie $(1,1)$ e $(-1,-1)$, $(1,-1)$ e $(-1,1)$ che ammettono come inverso se stesse. □

Problema: Mostrare che non tutti gli elementi di $Z\times(Z-\{0\})$ ammettono opposto. Pertanto $(Z\times(Z-\{0\}), +, \cdot, (0,1), (1,1))$ non è nemmeno un anello.

Dalla struttura $(Z\times(Z-\{0\}), +, \cdot, (0,1), (1,1))$ possiamo tuttavia ottenere un campo introducendo una opportuna congruenza e passando a quoziente. Ancora una volta possiamo ritenere equivalenti due coppie se "producono lo stesso effetto", allora, ad esempio, dobbiamo identificare la coppia (3,4) con la coppia (6,8). Per convincersi di questo fatto basta ricorrere alle solite torte che vengono proposte ai bambini a cui si insegnano le frazioni ed accorgersi che tre quarti di una torta coincidono con i sei ottavi di una torta. Ciò conduce a definire la seguente relazione.

Proposizione 9.2. Sia $\equiv$ la relazione in $Z\times(Z-\{0\})$ definita ponendo,

$$(p,q) \equiv (p',q') \Leftrightarrow p\cdot q' = q\cdot p'. \qquad (9.2)$$

Allora $\equiv$ è compatibile con le operazioni di somma e prodotto date in (10.1) ed è pertanto una congruenza.

Dim. Per provare che $\equiv$ è riflessiva basta osservare che $(p,q) \equiv (p,q)$ se e solo se $p\cdot q = q\cdot p$. Per provare che è simmetrica basta osservare che da $p\cdot q' = q\cdot p'$ segue che $p'\cdot q = q'\cdot p$. Per provare la proprietà transitiva osserviamo che

$$(n,m) \equiv (n',m'), (n',m') \equiv (p',q')$$
$$\Rightarrow n\cdot m' = m\cdot n' \,,\, n'\cdot q' = m'\cdot p'$$
$$\Rightarrow n\cdot m'\cdot n'\cdot q' = m\cdot n'\cdot m'\cdot p'$$
$$\Rightarrow n\cdot q'\cdot m\cdot p' \Rightarrow (n,m) \equiv (p',q').$$

Per provare la compatibilità con il prodotto osserviamo che

$$(n,m) \equiv (n',m'), (p,q) \equiv (p',q') \Rightarrow n\cdot m' = m\cdot n' \,,\, p\cdot q' = q\cdot p'$$
$$\Rightarrow n\cdot m'\cdot p\cdot q' = m\cdot n'.\, q\cdot p' \Leftrightarrow (n\cdot p,m\cdot q) \equiv (n'\cdot p',m'\cdot q')$$

$\Leftrightarrow (n,m)\cdot(p,q) \equiv (n',m')\cdot(p',q')$.

Similmente si prova la compatibilità rispetto la somma. □

Da notare che in termini di frazioni poniamo $p/q \equiv p'/q'$ se e solo se $p \cdot q' = q \cdot p'$. Ad esempio $6/8 \equiv 3/4$ in quanto $4\cdot 6 = 3\cdot 8$.

Teorema 9.3. La struttura $(Q, +, \cdot, 0, 1)$ ottenuta come quoziente modulo $\equiv$ di $(Z\times(Z\text{-}\{0\}), +, \cdot, (0,1), (1,1))$ è un campo che chiamiamo *campo dei numeri razionali*. Tale campo estende l'anello degli interi Z.

Dim. La struttura $(Q, +, \cdot, 0, 1)$ è definita dalle equazioni

$$[(p,q)] = \{(p',q') \mid (p',q') \equiv (p,q)\}$$

$$Q = \{[(p,q)] \mid (p,q) \in N_0 \times N_0,\ q \neq 0\}$$

$$[(p,q)]+[(a,b)] = [(p\cdot b+q\cdot a,\ q\cdot b)]$$

$$[(p,q)]\cdot[(a,b)] = [(p\cdot a,\ q\cdot b)].$$

$$0 = [(0,1)] \ ; \ 1 = [(1,1)].$$

E' immediato verificare che $[(0,1)]$ e $[(1,1)]$ sono elementi neutri rispetto all'addizione ed alla moltiplicazione, rispettivamente. Per provare che $[(p,q)]$ ammette opposto, osserviamo che $[(p,q)]+[(-p,q)] = [(p\cdot q\text{-}p\cdot q,q\cdot q)] = [(0,q\cdot q)]$. D'altra parte, poiché $(0,1) \equiv (0,q\cdot q)$, la classe $[(0,q\cdot q)]$ coincide con la classe $[(0,1)] = 1$. Quindi l'opposto di $[(p,q)]$ è $[(-p,q)]$. Ad esempio $[(3,4)]$ ammette come inverso $[(-3,4)]$ poiché $[(3,4)]+[(-3,4)] = [(12+-12,16)] = [(0,16)]$ e, poiché $(0,16) \equiv (0,1)$, $[(0,16)] = [(0,1)]$.

Sia $[(p,q)]$ non nullo e cioè tale che $p\neq 0$, allora $[(q,p)]$ è l'inverso di $[(p,q)]$. Infatti, essendo $(p\cdot q,q\cdot p)$ equivalente a $(1,1)$,

$$[(p,q)]\cdot[(q,p)] = [(p\cdot q,q\cdot p)] = [(1,1)].$$

Ad esempio $[(3,4)]$ ammette come inverso $[(4,3)]$ poiché $[(3,4)]\cdot[(4,3)] = [(12,12)]$ e, poiché $(12,12) \equiv (1,1)$. Le rimanenti proprietà di campo sono semplici da dimostrare.

Per mostrare che il campo $(Q, +, \cdot, 0, 1)$ è una estensione dell'anello degli interi relativi basta osservare che ponendo $f(n) = [(n,1)]$ si ottiene una immersione di Z in Q. □

Problema: Dire quale è l'inverso del numero razionale $[(2,3)]$ e dire il perché.

Problema: Considerando la definizione di campo dei razionali che abbiamo dato, dire che cosa significa che 1/2+3/6 = 1 ? Cosa significa che 10/-2 = -5 ?

Problema. Poiché siamo liberi di definire le operazioni in un insieme per inventare nuove strutture algebriche, definiamo nell'insieme Q dei numeri razionali l'operazione $\oplus$ ponendo

$$n/m \oplus p/q = (n^2+p^2)/(m^2+q^2).$$

1. Dire se l'operazione definita in questo modo è commutativa.
2. Dire perché è sbagliato chiedersi se l'operazione definita in questo modo è commutativa.

Concludiamo dicendo che nel campo dei razionali è definita una relazione d'ordine che lo rende un campo ordinato. Omettiamo la dimostrazione che si riduce ad una semplice verifica.

Proposizione 9.4. Definiamo nel campo dei razionali una relazione $\leq$ definita ponendo

$$[(p,q)] \leq [(n,m)] \Leftrightarrow p\cdot m \leq q\cdot n.$$

Allora il campo dei razionali con tale relazione diviene un campo ordinato.

Problema. Dire chi è maggiore tra i numeri 37/23 e 38/24.
Problema. Che cosa significa ridurre allo stesso denominatore.

Definizione 9.5. Chiamiamo *campo ordinato dei numeri razionali* il campo ordinato $(Q, +, \cdot, \leq, 0, 1)$.

10. Campo dei reali tramite le sezioni

Più delicato, da un punto di vista filosofico, è il passaggio dai razionali ai reali. Infatti in ogni definizione dei numeri reali viene necessariamente coinvolto in qualche modo l'infinito attuale.

Esponiamo, ad esempio, il metodo delle sezioni di Dedekind che è quello più utilizzato per la costruzione del campo dei reali anche se a mio parere è alquanto ferraginoso. Torniamo alla teoria delle grandezze omogenee che abbiamo esposto nel primo capitolo e supponiamo che in una classe $(G, =, <, +)$ di grandezze omogenee sia stata fissata una unità di misura $u \in G$. Allora, come abbiamo già osservato nel primo capitolo, se g è una grandezza da misurare un primo tentativo di misurazione consisterà nel prendere multipli successivi di u fino a raggiungere g. Se si trova un numero naturale m tale che $m \cdot u = g$ allora è possibile concludere che la misura di g rispetto ad u è m. Altrimenti si considera un naturale m tale che $m \cdot u < g < (m+1) \cdot u$ ed in tale caso si dice che m è una *misura per difetto* e $m+1$ una *misura per eccesso* di g. Una misurazione più precisa si può comunque ottenere dividendo u in q parti uguali (in generale in dieci parti) ed assumendo come sotto-unità di misura $u' = u/q$. Ora potrebbe capitare che per un opportuno m risulti che $m \cdot u' = (m/q) \cdot u = g$. In tale caso si concluderebbe che la misura cercata è m/q. Se invece ciò non accade allora potremmo lo stesso trovare m tale che $m \cdot u' < g < (m+1) \cdot u'$ e quindi $(m/q) \cdot u < g < (m+1)/q \cdot u$ e concludere che m/q è una misura per difetto e $(m+1)/q$ una misura per eccesso di g. Una misura più precisa si può avere dividendo la nuova unità di misura u' in un numero abbastanza alto di parti uguali. Nel caso in cui u e g siano incommensurabili, cioè che non esista un razionale m/q tale che $g = (m/q) \cdot u$ (come nel caso della diagonale e del lato del quadrato) tale processo di approssimazione non finisce mai. Allora in tale caso possiamo comunque considerare l'insieme A dei razionali positivi che misurano per difetto g e l'insieme B dei razionali positivi che misurano per eccesso g

$$A = \{p/q \in Q^+ \mid (p/q) \cdot u < g\}, \quad B = \{p/q \in Q^+ \mid (p/q) \cdot u > g\}.$$

E' facile verificare che:

a) $A \cap B = \emptyset$
b) $A \cup B = Q^+$
c) $x \in A, y \leq x \Rightarrow y \in A \quad ; \quad x \in B, y \geq x \Rightarrow y \in B.$
d) A è privo di massimo, B è privo di minimo.

Il metodo delle sezioni in un certo senso chiama "numero irrazionale positivo" una coppia di sottoinsiemi di Q^+ di questo tipo.[16]

---

[16]Ecco quanto afferma Dedekind a tale proposito:
Ora, in ogni caso in cui c'è una sezione (A, B) che non è prodotta da un numero razionale, allora noi creiamo un nuovo numero irrazionale che riteniamo completamente definito da questa sezione; diremo che questo numero corrisponde a questa sezione oppure che produce questa sezione.
Da notare l'espressione *"creiamo"* che testimonia che, almeno una parte della matematica per Dedekind è una creazione dell'uomo.

Definizione 10.1. Una *numero irrazionale positivo* è una coppia (A,B) di sottoinsiemi verificanti le condizioni *a*), *b*), *c*) e *d*).

Ora il nostro scopo è immergere sia i razionali che gli irrazionali in un unico ambiente in modo da potere parlare in generale di numero reale. A tale scopo possiamo identificare un numero razionale r con la coppia (A_r,B_r) di insiemi di razionali dove

$$A_r = \{x\in Q^+ \mid x<r\} \quad ; \quad B_r = \{x\in Q^+ \mid x>r\}.$$

In questo caso sono verificate le proprietà *a*), *c*) e *d*) mentre al posto della proprietà *b*) risulta che $A\cup B = Q^+ - \{r\}$. In altri termini se u e g sono commensurabili, ad esempio $g = r\cdot u$, per questione di uniformità di notazione indicheremo con (A_r, B_r) la misura di g rispetto ad u. In ogni caso viene individuata una coppia (A,B) di sottoinsiemi di Q^+ in cui, ripetiamo, la prima componente è vista come l'insieme delle misure per difetto e la seconda componente come l'insieme delle misure per eccesso di g.

Definizione 10.2. Chiamiamo *sezione positiva* o *numero reale positivo* una coppia (A, B) di insiemi di razionali tali che:

a) $A\cap B = \emptyset$

b) $A\cup B = Q^+$ oppure esiste $r\in Q^+$ tale che $A\cup B = Q^+ - \{r\}$

c) $x\in A\ , y\leq x \Rightarrow y\in A \quad ; \quad x\in B\ , y\geq x \ \Rightarrow\ y\in B.$

d) A è privo di massimo, B è privo di minimo.

Diremo che (A,B) è *irrazionale* se $A\cup B = Q^+$, che (A,B) è *razionale* se esiste $r\in Q^+$ tale che $A\cup B = Q^+ - \{r\}$.

Nel seguito indicheremo

- con 1 la sezione $(\{x\in Q^+ \mid x<1\},\{x\in Q^+ \mid x>1\})$
- con 0 la sezione $(\{x\in Q^+ \mid x<0\},\{x\in Q^+ \mid x>0\})$.

Più in generale, dato un razionale r indicheremo ancora con r la sezione (A_r, B_r) da esso determinata.

Un esempio di sezione che non è un razionale è il seguente

$$A = \{x\in Q^+ \mid x^2<2\} \quad ; \quad B = \{x\in Q^+ \mid x^2>2\}$$

che, in un certo senso, rappresenta il numero $\sqrt{2}$.

Per definire le operazioni aritmetiche tra numeri reali positivi, dati due insiemi X ed Y di numeri reali, poniamo

$$X+Y = \{x+y \mid x\in X, y\in Y\} \quad ; \quad X\cdot Y = \{x\cdot y \mid x\in X, y\in Y\}.$$

Definizioni 10.3. Chiamiamo *struttura algebrica dei reali positivi* la struttura $R^+ = (R^+, +, \cdot, 0, 1)$ dove

- R^+ è l'insieme delle sezioni positive
- l' operazione + è definita ponendo:

$$(A,B) + (A',B') = (A+A', B+B')$$

- l'operazione $\cdot$ è definita ponendo

$$(A,B)\cdot(A',B') = (A\cdot A', B\cdot B').$$

Inoltre si definisce un ordinamento $\leq$ ponendo

$$(A,B) \leq (A', B') \Leftrightarrow a \leq b' \text{ per ogni } a\in A \text{ e } b'\in B'.$$

Proposizione 10.4. La somma ed il prodotto di due sezioni è ancora una sezione.

Dim. Ci limitiamo a provare che $(A\cdot A', B\cdot B')$ verifica la condizione *a*). Infatti osserviamo che se $a\in A$ e $b\in B$, $a' \in A'$, $b'\in B'$, allora essendo il prodotto strettamente crescente (sui razionali positivi) e risultando che $a<b$ e $a' < b'$, possiamo asserire che $a\cdot a' < b\cdot b'$. Pertanto $(A\cdot A')\cap(B\cdot B') = \emptyset$. □

In tale modo si definisce la struttura algebrica dei reali positivi. Successivamente si procede a simmetrizzare tale struttura con un metodo analogo a quello che ha permesso di costruire l'anello *Z* degli interi relativi a partire dai numeri naturali. Non mi soffermo nei particolari di tale procedimento anche perché mi sembra preferibile definire i reali tramite il metodo che esporremo nel prossimo paragrafo.[17]

11. Campo dei reali tramite le successioni di Cauchy

Il metodo delle sezioni per definire i reali anche se è perfetto da un punto di vista formale, non corrisponde molto all'esperienza di chi si trova effettivamente a manipolare tali numeri. Infatti quando si utilizza un numero reale o lo si rappresenta come

[17] Usualmente nei libri di testo la nozione di sezione viene definita a partire dall'intero insieme *Q* dei razionali. Ciò permette di evitare il processo di simmetrizzazione. Ho preferito riferirmi solo alla parte positiva per mettere in evidenza la relazione tra la nozione di sezione ed il problema della misura. Inoltre il coinvolgimento dei razionali negativi renderebbe la definizione del prodotto poco naturale e noiosa per il fatto che il prodotto di due razionali negativi è un razionale positivo.

espansione decimale infinita (quindi come serie di potenze) oppure, più in generale, tramite una successione di razionali il cui limite è il numero reale in questione.

Una definizione del campo dei numeri reali che è molto più vicina a questo modo di procedere si ottiene utilizzando la nozione di successione di Cauchy. Essa utilizza metodi di analisi matematica ed è dovuta a Cantor il quale, non dimentichiamolo, ha sviluppato la teoria degli insiemi a partire da ricerche legate, appunto, all'analisi matematica.

Poiché nel campo dei razionali è definita una relazione d'ordine, possiamo parlare di razionale positivo e di razionale negativo. Inoltre possiamo definire la funzione *valore assoluto* ponendo $|x| = x$ se $x \geq 0$ e $|x| = -x$ altrimenti e quindi la funzione distanza $d : Q \to Q$ ponendo $d(x, y) = |x-y|$. Si ottiene in questo modo uno spazio metrico (Q, d) in cui, come è noto, sono definibili tutte le nozioni di analisi matematica come quella di successione convergente, di successione di Cauchy e così via.

Partiamo dalla potenza diretta $(Q^N, +, \cdot, \underline{0}, \underline{1})$ di Q con insieme di indici N. Tale struttura è definita assumendo che:
- il dominio sia l'insieme Q^N delle successioni di numeri razionali,
- l'addizione + sia definita ponendo

$$(a_n)_{n\in N} + (b_n)_{n\in N} = (a_n+b_n)_{n\in N}$$

- la moltiplicazione sia definita ponendo;

$$(a_n)_{n\in N} \cdot (b_n)_{n\in N} = (a_n \cdot b_n)_{n\in N}$$

- $\underline{0}$ denoti la successione $(z_n)_{n\in N}$ con $z_n = 0$ per ogni n
- $\underline{1}$ denoti la successione $(u_n)_{n\in N}$ con $u_n = 1$ per ogni n.

Proposizione 11.1. La struttura $(Q^N, +, \cdot, \underline{0}, \underline{1})$ è un anello unitario. Tuttavia, tale anello non è un campo.

Dim. [18] Per provare la proprietà commutativa della somma, che si esprime con una equazione del tipo $x+y = y+x$, osserviamo che

[18]Per provare che Q^N è un anello possiamo anche fare considerazioni di algebra universale. Infatti gli assiomi che caratterizzano l'essere un anello unitario sono tutti espressi tramite equazioni ed in algebra universale si prova che se una equazione vale per una famiglia di strutture allora vale anche per il prodotto diretto di tale famiglia. Pertanto ogni potenza diretta di un anello unitario è ancora un anello unitario. In particolare la potenza diretta Q^N è un anello unitario. Si noti che il provare che Q^N non è un campo mostra che la proprietà di essere campo, che

$$(a_n)_{n\in N}+(b_n)_{n\in N} = (a_n+b_n)_{n\in N} = (b_n+a_n)_{n\in N} = (a_n)_{n\in N}+(b_n)_{n\in N}.$$

Per provare che $\underline{0} = (z_n)_{n\in N}$ è l' elemento neutro rispetto all'addizione, osserviamo che

$$(a_n)_{n\in N} + (z_n)_{n\in N} = (a_n+z_n)_{n\in N} = (a_n)_{n\in N}.$$

Per provare che $\underline{1} = (u_n)_{n\in N}$ è l' elemento neutro rispetto alla moltiplicazione, osserviamo che

$$(a_n)_{n\in N} \cdot (u_n)_{n\in N} = (a_n \cdot z_n)_{n\in N} = (a_n)_{n\in N}.$$

Gli altri assiomi di teoria degli anelli si dimostrano in modo altrettanto banale. Per dimostrare che Q^N non è un campo consideriamo una successione $(a_n)_{n\in N}$ diversa dalla successione nulla ma che abbia un elemento uguale a zero, ad esempio $a_1 = 0$. Allora tale successione non può ammettere inverso in quanto comunque si scelga una successione $(b_n)_{n\in N}$ risulterà che $a_1 \cdot b_1 = 0$ e ciò comporta che $(a_n)_{n\in N} \cdot (b_n)_{n\in N}$ non può coincidere con $(u_n)_{n\in N}$.[19] □

Per ottenere un campo dobbiamo considerare prima una opportuna sottostruttura di $(Q^N, +, \cdot, \underline{0}, \underline{1})$ e poi un opportuno quoziente. Utilizzeremo la nozione di successione di Cauchy che chiunque abbia studiato un po' di analisi matematica ha già incontrato insieme al famoso teorema per cui una successione di reali è convergente se e solo se è una successione di Cauchy.

Definizione 11.2. Chiamiamo *successione di Cauchy* di razionali una successione $(r_n)_{n\in N}$ in Q^N tale che:

$$\forall\varepsilon>0\ \exists m\ \forall p\geq m \forall q\geq m\ |r_p-r_q|\leq\varepsilon.$$

Indichiamo con *Ch* l'insieme delle successioni di Cauchy.

Esempio. La successione $(1/n)_{n\in N}$ è una successione di Cauchy. Infatti, dato $\varepsilon > 0$ se fisso $m \geq 2/\varepsilon$ risulta che per ogni $p\geq m$ e $q\geq m$ $1/p\leq 1/m \leq \varepsilon/2$ e $1/q\leq 1/m \leq \varepsilon/2$. Ne segue che per ogni $p\geq m$ e $q\geq m$, $|1/p-1/q| \leq 1/p+1/q \leq \varepsilon$.

---

usualmente si rappresenta con l'asserzione $\forall x(x\neq 0 \rightarrow \exists y(x\cdot y = 1)$ non può essere espressa da una equazione.

[19] Come osservato nell'appendice, un modo per provare che un anello non è un campo è provare che ammette divisori dello zero. Nel nostro caso la successione $(a_n)_{n\in N}$ è un divisore dello zero. Infatti, consideriamo la successione $(b_n)_{n\in N}$ con $b_1 = 1$ e $b_n = 0$ per ogni $n\neq 1$. Allora di $(a_n)_{n\in N}$ per $(b_n)_{n\in N}$ è la successione nulla anche se entrambe le successioni sono diverse dalla successione nulla.

Esempio. Dato un razionale r, indichiamo con $\underline{r}$ la successione costantemente uguale ad r, allora poiché $|r_p - r_q|$ è uguale a zero, $\underline{r}$ è una successione di Cauchy. In particolare sono successioni di Cauchy le successioni $\underline{0}$ ed $\underline{1}$.

In analisi matematica viene provata la seguente proposizione.

Proposizione 11.3. L'insieme *Ch* delle successioni di Cauchy è chiuso rispetto alla somma, al prodotto ed all'opposto. Pertanto la struttura $(Ch, +, \cdot, \underline{0}, \underline{1})$ è un sottoanello di $(Q^N, +, \cdot, \underline{0}, \underline{1})$**.** Tale anello non è un campo.

Dim. Ci limitiamo a provare che $(Ch, +, \cdot, \underline{0}, \underline{1})$ non è un campo procedendo come nella Proposizione 12.1. Infatti consideriamo la successione $(a_n)_{n\in N}$ definita ponendo $a_1 = 0$ ed $a_n = 1$ per ogni $n\neq 1$, allora è facile provare che $(a_n)_{n\in N}$ è una successione di Cauchy. Tale successione è diversa dalla successione nulla e non ammette inverso poiché data una qualunque successione di Cauchy $(b_n)_{n\in N}$ il prodotto $(a_n \cdot b_n)_{n\in N}$ non può essere la successione unitaria avendo il primo elemento $a_1 \cdot b_1 = 0 \cdot b_1 = 0 \neq 1$.

Problema 1. Esistono elementi invertibili in $(Q^N, +, \cdot, \underline{0}, \underline{1})$? Dimostrare un teorema del tipo "Un elemento di Q^N è invertibile se e solo se …".

Un po' più difficile è il seguente problema.

Problema 2. Esistono elementi invertibili nella struttura $(Ch, +, \cdot, \underline{0}, \underline{1})$? Dimostrare un teorema del tipo "Un elemento di *Ch* è invertibile se e solo se …". (Si suggerisce di analizzare se $(1/n)_{n\in N}$ è invertibile e se $(n/(1+2n))_{n\in N}$ è invertibile).

12. Passando a quoziente: il campo dei reali

Abbiamo visto che $(Q^N, +, \cdot, \underline{0}, \underline{1})$ non è un campo anche se esistono elementi di *Ch* che ammettono inverso. Ad esempio la successione $(n/(n+1))_{n\in N}$ ammette come inverso la successione $((n+1)/n)_{n\in N}$. D'altra parte chi ha studiato un po' di analisi matematica sa che il limite di una successione convergente di reali non cambia se noi alteriamo un numero finito di valori della successione. In altre parole se due successioni differiscono per un numero finito di valori il loro limite è lo stesso. Chiamiamo *quasi-ovunque uguali* due successioni che differiscono per un numero finito di valori. Allora anche il controesempio della Proposizione 12.3 non sarebbe più valido se noi cambiassimo il suo primo valore da 0 ad 1 e quindi passassimo dalla successione $(a_n)_{n\in N}$ alla successione costantemente uguale ad 1. Più in generale

tutte le successioni di Cauchy che hanno solo un numero finito di 0 sono quasi-ovunque uguali ad una successione di Cauchy che ammette inverso.

Per ottenere un campo dobbiamo considerare un opportuno quoziente della struttura $(Ch, +, \cdot, \underline{0}, \underline{1})$.

Definizione 12.1. Diciamo che due successioni di Cauchy $(a_n)_{n\in N}$ e $(b_n)_{n\in N}$ sono *equiconvergenti* e poniamo $(a_n)_{n\in N} \equiv (b_n)_{n\in N}$ se $lim_{n\to\infty} |a_n-b_n| = 0$.

Allora due successioni $(a_n)_{n\in N}$ e $(b_n)_{n\in N}$ sono equiconvergenti se

$$\forall \varepsilon>0\ \exists m\ \forall p\geq m\ |a_p-b_p| \leq \varepsilon.$$

Ad esempio la successione $(n/(n+1))_{n\in N}$ e la successione $\underline{1} = (u_n)_{n\in N}$ sono equiconvergenti poiché

$$|n/(n+1) - u_n| = |n/(n+1) -1| = |(n-n-1)/(n+1)| = 1/n.$$

la loro differenza è equiconvergente. Più in generale sono equiconvergenti a $\underline{1}$ tutte e sole le successioni convergenti ad 1. Similmente sono equiconvergenti a $\underline{0}$ tutte e sole le successioni convergenti a zero.

Proposizione 12.2. La relazione di equiconvergenza è una congruenza nell'anello $(Ch, +, \cdot , \underline{0} , \underline{1})$ e pertanto ha senso considerare il quoziente $(Ch/\equiv, + , \cdot , [\underline{0}], [\underline{1}])$ di tale anello.[20]

Dim. E' evidente che $\equiv$ è una relazione di equivalenza. Per provare che è compatibile con il prodotto, supponiamo che $(a_n)_{n\in N}$ sia equiconvergente a $(\underline{a}_n)_{n\in N}$ e che $(b_n)_{n\in N}$ sia equiconvergente a $(\underline{b}_n)_{n\in N}$. Dobbiamo provare che $(a_n\cdot b_n)_{n\in N}$ è equiconvergente a

[20] Chi conosce un po' di teoria degli anelli sa che definire una nozione di congruenza equivale a definire una nozione di ideale. Nel nostro caso invece di riferirci alla relazione di equiconvergenza possiamo riferirci all'ideale costituito dall'insieme I delle successioni di Cauchy $(a_n)_{n\in N}$ tali che $lim_{n\to}|a_n| = 0$. E' immediato che due successioni $(a_n)_{n\in N}$ e $(b_n)_{n\in N}$ sono equiconvergenti se e solo se la loro differenza appartiene ad I.

$(\underline{a}_n \cdot \underline{b}_n)_{n\in N}$. Posto $\delta_n = \underline{a}_n - a_n$ e $\gamma_n = \underline{b}_n - b_n$, per ipotesi le successioni $(\delta_n)_{n\in N}$ e $(\gamma_n)_{n\in N}$ convergono a zero. Inoltre, poichè

$$\begin{aligned} \underline{a}_n \cdot \underline{b}_n - a_n b_n &= (a_n+\delta_n)\cdot(b_n+\gamma_n) - a_n b_n \\ &= a_n\gamma_n + \delta_n \cdot b_n + \delta_n \gamma_n \end{aligned}$$

e poiché $(a_n \gamma_n)_{n\in N}$, $(\delta_n \cdot b_n)_{n\in N}$ e $(\delta_n \gamma_n)_{n\in N}$ convergono a zero, $(a_n \cdot b_n)_{n\in N}$ risulta equiconvergente a $(\underline{a}_n \cdot \underline{b}_n)_{n\in N}$. Similmente si prova che l'addizione è compatibile con l'equiconvergenza. □

La dimostrazione del seguente lemma è evidente.

Lemma 12.3. Siano $(a_n)_{n\in N}$ e $(b_n)_{n\in N}$ due successioni che coincidono quasi-ovunque, cioè per le quali esiste $m\in N$ tale che $a_n = b_n$ per ogni $n\geq m$, allora tali successioni sono equiconvergenti.

Ad esempio ne segue che una successione che assume quasi-ovunque il valore 5/4 è equiconvergente alla successione costantemente uguale a 5/4.

Lemma 12.4. Se una successione di Cauchy $(a_n)_{n\in N}$ assume infinite volte il valore r allora converge ad r. Pertanto una successione di Cauchy che non converge a zero può assumere al più un numero finito di volte il valore 0.

Dim. Sappiamo che, dato $\varepsilon > 0$ esiste m tale che, per ogni p e q maggiori di m, $|a_p - a_q| \leq \varepsilon$. Allora poiché esistono infiniti indici in cui la successione assume il valore r ne esiste anche uno $\underline{n}$ maggiore di m e quindi per tale numero risulterebbe che $|a_p - r| = |a_p - a_{\underline{n}}| \leq \varepsilon$ per ogni $p\geq m$. Ciò mostra che il limite di $(a_n)_{n\in N}$ è r.
Supponiamo che $(a_n)_{n\in N}$ sia una successione di Chauchy che non converge a zero e sia $(\underline{a}_n)_{n\in N}$ tale che $\underline{a}_n = 1$ se $a_n = 0$ e $\underline{a}_n = a_n$ altrimenti. Allora poiché le due successioni coincidono quasi ovunque, per il lemma 13.3 sono equiconvergenti. □

Non è detto che l'inverso di una successione di Cauchy sia ancora una successione di Cauchy. Ad esempio la successione $(1/n^2)_{n\in N}$ ha come inversa la successione $(n^2)_{n\in N}$ che non lo è. Tuttavia si prova le seguente proposizione.

Lemma 12.5. Se $(a_n)_{n\in N}$ è una successione di Cauchy che non converge a zero ed i cui elementi sono non nulli, allora $(1/a_n)_{n\in N}$ è ancora una successione di Cauchy.

Teorema 12.6. Il quoziente $(Ch/\equiv, +, \cdot, [\underline{0}], [\underline{1}])$ è un campo che estende il campo dei numeri razionali. In tale campo esiste x tale che $x^2 = 2$.

Dim. Poiché il quoziente di un anello unitario è ancora un anello unitario, per provare che $(Ch/\equiv, +, \cdot, [\underline{0}], [\underline{1}])$ è un campo dobbiamo solo provare che l'insieme degli elementi non nulli costituisce un gruppo rispetto alla moltiplicazione. In altre parole, dobbiamo provare che

$$[(a_n)_{n\in N}] \neq [\underline{0}] \Rightarrow \exists (b_n)_{n\in} \in Ch, [(a_n)_{n\in N}]\cdot[(b_n)_{n\in N}] = [\underline{1}].$$

Ora poiché $[(a_n)_{n\in N}] \neq [\underline{0}]$ per il Lemma 13.4 non è restrittivo supporre che $(a_n)_{n\in N}$ sia stato scelto in modo che non assuma mai valore zero. Allora per quanto detto nel lemma 13.5, $(1/a_n)_{n\in N}$ è ancora una successione di Cauchy. E' evidente che $[(a_n)_{n\in N}]\cdot[(1/a_n)_{n\in N}] = [\underline{1}]$.

Per provare che $(Ch/\equiv, +, \cdot, [\underline{0}], [\underline{1}])$ estende il campo dei razionali basta osservare che la funzione $f: Q \rightarrow Ch/\equiv$ definita ponendo $f(r) = [\underline{r}]$ è una immersione.

Infine se $(q_n)_{n\in N}$ è una successione di razionali tale che $lim_{n\rightarrow\infty} q_n^2 = 2$, allora il numero reale $r = [(q_n)_{n\in N}]$ ha come quadrato 2. Ad esempio l'usuale regola per l'estrazione della radice di un numero è in grado di fornire una tale successione. □

Per definire l'ordine procediamo al modo seguente.

Definizione 12.7. In $Ch/\equiv$ definiamo la relazione $<$ ponendo $a<a'$ se esistono due successioni di Cauchy $(a_n)_{n\in N}$ e $(b_n)_{n\in N}$ tali che $a = [(a_n)_{n\in N}]$ e $b = [(b_n)_{n\in N}]$ per cui esiste m tale che $a_n < b_n$ per ogni $n \geq m$.

Questa relazione è di ordine stretto e, come al solito determina una relazione d'ordine $\leq$ ponendo $[(a_n)_{n\in N}] \leq [(b_n)_{n\in N}]$ se $[(a_n)_{n\in N}] < [(b_n)_{n\in N}]$ oppure $[(a_n)_{n\in N}] = [(b_n)_{n\in N}]$.

Problema. Perché abbiamo definito prima l'ordine stretto o poi l'ordine?

Omettiamo la dimostrazione del seguente fondamentale teorema.

Teorema 12.8. La struttura ($Ch/\equiv, +, \cdot, \leq, [\underline{0}], [\underline{1}]$) è un campo ordinato completo[21] che chiamiamo *campo dei numeri reali.*

13. Metodo delle successioni "nidificate" di intervalli

Questo metodo è in qualche modo legato alla *"interval analisis"* cioè a un metodo di calcolo che si basa sugli intervalli chiusi visti come rappresentazione approssimativa dei numeri reali. Cominciamo con la seguente definizione.

Definizione 13.1. Indichiamo con *IR* l'insieme degli intervalli non vuoti, chiusi e limitati del campo ordinato dei razionali. Chiamiamo *imprecisione* di un intervallo $[a, b]$ in *IR* la differenza $b-a$.

Interpretiamo un elemento $[a, b]$ di *IR* come un'informazione su di un valore di una grandezza che per qualche motivo non riusciamo ad individuare perfettamente. Tale informazione ci dice che *"il numero si colloca tra a e b"*, e potrebbe derivare da una misurazione di una grandezza di tipo fisico o dal troncamento di qualche algoritmo quando si ritiene che l'accuratezza del risultato sia sufficientemente preciso per i calcoli che si devono fare. Se l'imprecisione è nulla allora l'informazione è completa ed individua il numero razionale $q = a = b$.

Vogliamo poi definire in *IR* alcune operazioni che, come vedremo, estendono a *IR* le usuali operazioni in *Q*.

Definizione 13.2. Possiamo estendere le due operazioni + e · del campo *Q* alla classe *IR* ponendo

$$A+B = \{a+b : a\in A, b\in B\} \quad ; \quad A\cdot B = \{a\cdot b : a\in A, b\in B\}$$

per ogni *A* e *B* in *IR*. Inoltre possiamo identificare 0 ed 1 con gli intervalli $[0, 0] = \{0\}$ e $[1, 1] = \{1\}$.

Si dimostra che se *A* e *B* sono intervalli chiusi e limitati anche $A+B$ e $A\cdot B$ sono intervalli chiusi e limitati e quindi che le operazioni ora definite sono interne a *IR*. Quindi determinano una

[21] Per la definizione di tale nozione veda l'Appendice.

struttura algebrica (IR, +, ·, 0, 1) dove 0, 1 denotano gli intervalli [0, 0] e [1,1], rispettivamente. Tali definizioni si giustificano dal fatto che se i due numeri a e b cadono negli intervalli A e B, allora $a+b$ cade sicuramente in $A+B$. Questo significa che $A+B$ è l'informazione su $a+b$ che possiamo ricavare dall'informazioni A e B. Lo stesso si può dire del prodotto.

Proposizione 13.3. La funzione $i : Q \to IR$ definita ponendo $i(x) = [x, x] = \{x\}$ è una immersione di (Q, +, ·, 0, 1) in (IR, +, ·, 0, 1).

Esistono formule che facilitano i calcoli. Infatti abbiamo che, per ogni coppia $[a, b]$ e $[a', b']$ in IR,

$[a, b]+[a', b'] = [a+a', b+b']$;
$[a,b]\cdot[a',b']=[min\{a\cdot a',a\cdot b',b\cdot a',b\cdot b'\},$
$max\{a\cdot a',a\cdot b',b\cdot a',b\cdot b'\}]$;

La definizione apparentemente macchinosa della moltiplicazione è dovuta alla regola dei segni. Ecco alcuni esempi:

$[0, 1]+[2, 3] = [2,4]$, $[1,2]-[1, 2] = [-1, 1]$, $[-3, 2]\cdot[1, 2] = [-6, 4]$, $[1, 2]/[3, 4] = [1/4, 2/3]$.

Da notare che possiamo definire in IR anche le due operazioni corrispondenti alla sottrazione – e la divisione / in Q. Tuttavia si deve considerare A/B non definito nel caso in cui $0\in B$. E facile anche vedere che
$[a, b]-[a', b'] = [a-b', a'-b]$ e
$[a,b]/[a',b']=[min\{a/a',a/b',b/a',b/b'\},$
$max\{a/a',a/b',b/a',b/b'\}]$.
Una relazione binaria < in IR si può definire ponendo

$A < B \Leftrightarrow$ per ogni $a\in A$ e $b\in B$ risulta $a<b$.

La giustificazione per tale definizione è che se sappiamo che $A < B$ e che due numeri reali sono in A e B, allora siamo anche sicuri che il primo è minore del secondo. Una tale definizione può essere anche riscritta nella forma

$[a, b] < [a', b'] \Leftrightarrow b < a'$.

In questo modo si definisce una nuova struttura (IR, +, -, ·, /, <) che noi indicheremo semplicemente con IR.[22] IR estende il campo ordinato dei numeri razionali poiché possiamo identificare ogni razionale q con l'intervallo $[q,q]$. Precisamente la funzione $i : Q \rightarrow IR$ che abbiamo prima definito è una immersione di (Q, +, -, ·, /, <) in (IR, +, -, ·, /, <) e quindi è un isomorfismo di Q nella sottostruttura $i(Q)$ di IR. Tuttavia IR non è un campo. Infatti pur essendo l'addizione e la moltiplicazioni operazioni commutative, e pur essendo [0, 0] ed [1, 1] elementi neutri rispetto tali operazioni, risulta che dato $[a, b]$ con a diverso da b $[a, b]+[x, y] = [a+x, b+y] = [0, 0]$ solo se $y+x = 0$ e $b+y = 0$ e quindi solo se se solo se $x = -a$ e $y = -b$. Ma questo è assurdo in quanto essendo $b-a>0$ dovrebbe risultare che $y-x = a-b <0$.

Passiamo ora alla definizione di numero reale.

Definizione 13.4. Una *successione nidificata di intervalli razionali*, è una successione $([a_n, b_n])_{n\in N}$ decrescente di elementi di IR tale che $lim_{n\rightarrow\infty} b_n - a_n = 0$. Indichiamo con SN l'insieme di tali successioni e con (SN, +, -, ·, /, ≤) la struttura algebrica definita in SN ponendo, per ogni operazione $*$

$$([a_n, b_n])_{n\in N} * ([a'_n, b'_n])_{n\in N} = ([a_n, b_n] * [a'_n, b'_n])_{n\in N}.$$

Una successione nidificata rappresenta una successione di approssimazioni di un numero reale sempre più precise in cui l'errore b_n-a_n diventa nullo per n che tende all'infinito. Ora $([a_n, b_n])_{n\in N}$ e $([a'_n, b'_n])_{n\in N}$ definiscono sicuramente due numeri reali diversi solo se esiste $n\in N$ tale che $[a_n, b_n]\cap[a'_n, b'_n] = \emptyset$, cioè se le informazioni rappresentate da tali intervalli diventano inconsistenti tra loro al passo n. Questo suggerisce la seguente definizione.

Proposizione 13.5. La relazione ≡ definite ponendo $([a_n, b_n])_{n\in N} \equiv ([a'_n, b'_n])_{n\in N}$ se $[a_n, b_n]$ si sovrappone a $[a'_n, b'_n]$ per ogni $n\in N$ è una congruenza.

---

[22] Tale struttura è alla base di un interessante approccio al calcolo numerico noto sotto il nome di *interval analysis*. Tale teoria infatti fornisce uno strumento capace di dominare l'ampiezza degli errori che si verificano nei calcoli.

Definizione 13.3. Chiamiamo *insieme dei numeri reali* il quoziente $SN/\equiv$ della struttura $(SN, +, \cdot, \leq)$ modulo $\equiv$.

14. Metodo dei filtri di Cauchy

In realtà il problema di definire i punti in una struttura point-free è da lungo tempo considerato dai matematici che si occupano della teoria dei reticoli o delle algebre di Boole. Il contesto non è geometrico e riguarda teoremi di rappresentazione di tale tipo di strutture. In questo ambito un ruolo fondamentale viene giocato dai filtri primi i quali vengono proposti come una sorta di definizione di punto. Ricordiamo la definizione di filtro.

Definizione 14.1. Dato un insieme non vuoto S chiamiamo *filtro* su S una classe F di sottoinsiemi di S tale che:

$X\in F,\ Y\in F \Rightarrow X\cap Y\in F$

$X\in F,\ Y\supseteq X \Rightarrow Y\in F.$

Un filtro si dice *primo* se inoltre

$X\cup Y\in F \Rightarrow X\in F$ oppure $Y\in F.$

Un esempio di filtro primo in un insieme S si ottiene fissando un elemento P di S e ponendo $F = \{X\in P(S) : P\in X\}$. Un esempio di filtro non primo è il filtro dei sottoinsiemi co-finiti di S, cioè dei sottoinsiemi il cui complemento è finito o vuoto. Nel piano Euclideo un esempio è dato dalla classe degli intorni di un punto.

Vogliamo mostrare come sia possibile utilizzare tale nozione per dare una possibile definizione del campo dei reali. In questo caso dobbiamo considerare un particolare tipo di filtri su Q.

Definizione 14.2. Chiamiamo *filtro di Cauchy* un filtro F in Q tale che per ogni numero naturale $n\in N$ esiste in F un intervallo $[a, b]$ la cui ampiezza $b\text{-}a$ è minore di $1/n$.

La seguente proposizione mostra che tale nozione è strettamente collegata con quella delle successioni nidificate di intervalli razionali.

Proposizione 14.3. Un filtro su Q è un filtro di Cauchy se e solo se contiene una successione nidificata di intervalli in Q.

Dim. E' evidente che se un filtro F contiene una successione nidificata di intervalli in Q allora è un filtro di Cauchy. Vice-versa, sia F un filtro di Cauchy e sia, per ogni $n\in N$, $[\underline{a}_n, \underline{b}_n]$ un intervallo in F il cui diametro è minore di $1/n$. Allora la successione $([a_n, b_n])_{n\in N}$ definita per ricorsione tramite le equazioni

$[a_1, b_1] = [\underline{a}_1, \underline{b}_1]$

$[a_n, b_n] = [a_{n-1}, b_{n-1}]\cap[\underline{a}_n, \underline{b}_n]$ se $n \neq 1$

è una successione nidificata di intervalli in Q i cui elementi appartengono ad F. Inoltre tutti gli insiemi in F contengono un intervallo $[\underline{a}_n, \underline{b}_n]$ e quindi contengono l'intervallo $[a_n, b_n]$.

Questo significa che possiamo definire i filtri di Cauchy come i filtri generati da una successione nidificata di intervalli.

Definizione 14.4. Chiamiamo *equivalenti* due filtri di Cauchy generati da due successioni equivalenti di intervalli e *numero reale* ogni elemento del quoziente CF/≡ di CF modulo tale equivalenza.

Concludiamo qui questa esposizione poiché il libro è indirizzato principalmente al mondo della scuola e non mi sembra che la nozione di filtro possa essere utile in tale contesto. Chi vuole approfondire questo tipo di approccio può trovare una buona trattazione in "Weiss I., *The reals as rational Cauchy filters*, arXiv:1503.04348v3 [math.HO] 5 Nov 2015."

15. La rappresentazione dei numeri

Tutti sappiamo che con una espressione come "3507" viene indicato il numero naturale che è la somma di tre migliaia, cinque centinaia, zero decine e sette unità. In questo caso si dice anche che "3507" è la rappresentazione in base dieci di tale numero. Una rappresentazione di questo tipo utilizza simboli presi nell'insieme $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, simboli che sono messi uno dopo l'altro. Nella teoria dei linguaggi formali, che vedremo nel secondo volume, si dice che A è un *alfabeto* e che "3507" è una *parola* in tale alfabeto.

Definizione 15.1. Un *alfabeto* è un insieme finito i cui elementi vengono detti *lettere*, una *parola su tale alfabeto* è una sequenza $a_m\ldots a_0$ di lettere e si indica con A^+ l'insieme delle parole sull'alfabeto A. Una *rappresentazione* di un insieme S in A^+ è una funzione iniettiva $c : S \to A^+$.

Abbiamo usato l'espressione "sequenza" e non l'espressione "insieme" poiché l'ordine con cui sono indicati gli elementi di A in una parola e la presenza di eventuali ripetizioni delle lettere sono importanti. Le parole 345, 543, 3445 sono da intendere diverse tra loro pur essendo formate con gli stessi elementi di A.

Tornando alla rappresentazione in base 10, abbiamo visto che ogni parola nel linguaggio $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ (che non inizi con 0) può essere interpretata come un numero naturale. Si pone allora subito il problema se in questo modo tutti i numeri naturali siano rappresentabili tramite una parola in A. Fortunatamente vale il seguente teorema.

Teorema 15.2. Sia b un intero diverso da 0 ed 1, allora per ogni numero naturale n esistono $a_m,\ldots,a_0$ in $\mathbb{N}_0$ minori di b tali che

$$n = a_m \cdot b^m + \ldots + a_0 \cdot b^0.$$

Se $n \neq 0$ e poniamo la condizione $a_m \neq 0$, tali interi sono unici.

Dim. Vogliamo utilizzare il principio di induzione transfinita. A tale scopo, fissato n, supponiamo che il teorema valga per tutti i numeri $x<n$. Poiché per l'algoritmo della divisione $n = b \cdot c + r$ con $c<n$ il teorema è vero per c e quindi esistono interi $a_m,\ldots,a_0$ minori di b tali che

$$n = b \cdot (a_m \cdot b^m + \ldots + a_0 \cdot b^0) + r = a_m \cdot b^{m+1} + \ldots + a_0 \cdot b^1 + r \cdot b^0.$$

Ne segue che il teorema è vero anche per n.
L'unicità dei coefficienti $a_m,\ldots,a_0$ segue dal fatto che ogni a_i si ottiene come resto di una divisione e che il resto di una divisione è determinato univocamente. □

Se allora rappresentiamo con un opportuno insieme A di b simboli i primi b numeri in $\mathbb{N}_0$, possiamo rappresentare ogni numero n con una parola in A^*.

Definizione 15.3. Sia $b \neq 1$, $b \neq 0$, sia A un insieme di b simboli che chiamiamo *cifre* ed indichiamo ogni naturale a tale che $0 \leq a < b$ con una corrispondente cifra $\underline{a} \in A$. In tale ipotesi la *rappresentazione in base b* di un naturale n è la parola $\underline{a}_m \ldots \underline{a}_0$ nei simboli in A dove $n = a_m \cdot b^m + \ldots + a_0 \cdot b^0$ con $a_m \neq 0$ ed $0 \leq a_i < b$.

Nel caso in cui b è uguale a 10si ottiene l'usuale rappresentazione in base 10. Nel caso in cui b sia uguale a 2 possiamo utilizzare le

cifre 0 ed 1. Ad esempio in base due la parola 10010, indica il numero $1\cdot 2^4+0\cdot 2^3+0\cdot 2^2+1\cdot 2^1+0\cdot 2^0$, cioè il numero composto da
0 unità,
1 gruppo di 2,
0 gruppi di 4,
0 gruppi di 8
1 gruppo di 16
che corrisponde al numero in base 10 indichiamo con 18.
Se vogliamo rappresentare i numeri in base dodici, dobbiamo inventarci nuovi simboli per denotare dieci ed undici e quindi, ad esempio, porre $A = \{0, 1,2, 3, 4, 5, 6, 7, 8, 9, d, u\}$. In questo caso 23 denota il numero composto da due dozzine più 3 unità (ventisette), 2d denota il numero costituito da due dozzine più dieci unità (trentaquattro), dd è il numero costituito da dieci dozzine più dieci unità (centotrenta) e così via.

16. Rappresentare altri insiemi numerici
Per rappresentare i numeri relativi ed i razionali si deve tenere conto che tali insiemi si ottengono come quoziente $S/\equiv$ di un'altra struttura S. Questo suggerisce l'idea che si possano utilizzare una rappresentazione $c : S \to A^*$ già definita per S. Possiamo infatti procedere al modo seguente.

1. Si sceglie in ogni classe $[x]$ un elemento che si considera rappresentativo dell'intera classe. Questo significa avere una funzione di scelta $r : S/\equiv \to S$ tale che $r([x]) \in [x]$ o, equivalentemente, $r([x]) \equiv x$. Se indichiamo con S^* il codominio di r, tale funzione, evidentemente iniettiva, risulta biettiva come funzione di $S/\equiv$ in S^*.
2. Detta $c : S/\equiv \to A^*$ una rappresentazione di $S/\equiv$ nell'alfabeto A, assumiamo che la rappresentazione della classe $[x]$ sia quella dell'elemento $r([x])$ di S^* cioè poniamo $c([x]) = c(r([x])$.
2. Si definisce la funzione $\underline{r}$: $S \to S$ col porre $\underline{r}(x) = r([x])$, in questo caso si dice che $\underline{r}(x)$ *è il ridotto di* x e chiamiamo *"ridotti"* gli elementi di S^*. Poiché $r : S/\equiv \to S$ è iniettiva, la funzione $r : S/\equiv\to S^*$ risulta essere biettiva.

Facciamo alcuni esempi.

1. Interi relativi. Gli interi relativi li abbiamo definiti quozientando l'insieme $N_0\times N_0$ delle coppie di numeri naturali tramite la relazione per cui $(n, m) \equiv (n', m')$ se $n+m = n'+m'$. Questo ci ha consentito di scegliere in ogni classe $[(n, m)]$ un elemento equivalente che ha o primo elemento o econdo elemento uguale a zero. Quindi $r([n, m]) = (n\text{-}m, 0)$ se $n \geq 0$ e $r([n, m]) = (0, m\text{-}n)$ se $n >$. Se arricchiamo l'alfabeto con cui abbiamo rappresentato i numeri naturali aggiungendo i simboli + e -, allora possiamo associare alla parola riscrivere una coppia $(n, 0)$ nella parola $+n$ ed una coppia $(0, m)$ nella parola $–m$. In definitiva possiamo rappresentare la classe $[(n, m)]$ con la parola $+c(n\text{-}m)$ se $n \geq m$ e con la parola $\text{-}c(m\text{-}n)$ altrimenti. Ad esempio la classe [(7,2)] sarà rappresentata dalla parola +5. La classe [(2,6)] sarà rappresentata dalla parola –4. In questo modo si ottiene la nota rappresentazione dei numeri relativi come *numero+segno*.

2. Razionali. Nel caso dei razionali sappiamo che (x, y) è equivalente a $(x/d, y/d)$ dove d è il massimo comun divisore tra x ed y). Quindi conviene definire r ponendo $r([(x, y)]) = (x/d, y/d)$. In questo caso A^* è l'insieme delle coppie Tenendo conto che si usa utilizzare l'espressione p/q al posto di (p,q), un numero razionale viene rappresentato da espressioni del tipo del tipo p/q con p e q primi tra loro e $q>0$.

e se x ed y sono due elementi di A^* allora $x\underline{+}y = \underline{r}(x+y)$ e $x\underline{\cdot}y = \underline{r}(x \cdot y)$. Ad esempio poniamo $3\underline{+}4 = 2$, $3\underline{+}1 = 4$, $3\underline{\cdot}4 = 2$ e così via. Infine la funzione r è tale che

$$\begin{aligned} r([x]+[y]) &= r([\underline{r}(x)]+[\underline{r}(y)]) = r([\underline{r}(x)+\underline{r}(y)]) = \underline{r}(\underline{r}(x)+\underline{r}(y)) \\ &= \underline{r}(r([x])+r([y])) = r([x])\underline{+}r([y]). \end{aligned}$$

Similmente si prova che $r([x]\cdot[y]) = r([x])\underline{\cdot}r([y])$ e quindi r è un isomorfismo di $(Z/m, +, \cdot)$ in $(A^*, \underline{+}, \cdot)$.

3. **Interi modulo *m*.** Sia $Z/5$ l'anello degli interi modulo 5 e scegliamo in ogni classe $[x]$ il resto $r([x])$ della divisione di x per 5. Pertanto A^* coincide con l'insieme dei numeri da zero a quattro. Poiché tali numeri sono rappresentati, nell'ordine, dai simboli 0, 1, 2, 3, 4, abbiamo che le cinque classi di $Z/5$ sono rappresentate ancora da tali simboli.

Ovviamente r risulta essere una funzione biettiva di A/m in A^* e questo consente di *"proiettare"* la struttura quoziente $A/\equiv$ in A^*

al modo seguente. Precisamente, vogliamo definire in A^* operazioni in modo che r risulti essere un isomorfismo di $A/\equiv$ in A^*. Ad esempio vogliamo che se $*$ è un'operazione binaria in $A/\equiv$ allora esista una corrispondente operazione $\underline{*}$ in A^* per cui $r([x]*[y]) = r([x])\underline{*}r[y]$, cioè per cui $\underline{r}([x*y]) = \underline{r}(x)\ \underline{*}\ \underline{r}(y)$. Per ottenere questo basta definire $*$ ponendo $x\underline{*}y = \underline{r}(x*y)$. Allora $r(x')\underline{*}r(y') = \underline{r}(r(x')*r(y'))$
Detto in modo più formale, possiamo procedere nel modo seguente.

Definizione 16.1 Data una struttura algebrica A ed una congruenza $\equiv$ in A, chiamiamo *funzione di scelta* una funzione iniettiva $n : A/\equiv\ \rightarrow A$ tale che $n([x])\in[x]$. Chiamiamo *in forma normale* i valori di tale funzione, cioè gli elementi di $A^*=\{n([\{x\}]) : x\in A\}$. Poniamo poi $\underline{n}(x) = n([\{x\}])$.

Ovviamente $\underline{n}(x)$ è l'elemento in forma normale $\equiv$-equivalente a x. Spesso si dice che $\underline{n}(x)$ è la riduzione in forma normale di x.

Definizione 16.2. Data una struttura A, una congruenza $\equiv$ in A ed una funzione di scelta n, indichiamo con A^* la struttura il cui dominio è l'insieme A^* degli elementi di A in forma normale e le cui operazioni si ottengono trasformando ogni operazione n-aria h: $A^n\rightarrow A$ nell'operazione h^*: $(A^*)^n\rightarrow A^*$ ottenuta ponendo

$$h^*(x_1,\ldots,x_n) = n([h(x_1,\ldots,x_n)])$$
$$x*y = n(x*y).$$

Teorema 16.3. Data una struttura algebrica A ed una congruenza $\equiv$ in A, la funzione $n : A/\equiv\ \rightarrow A^*$ è un isomorfismo tra la struttura $A/\equiv$ e la struttura A^*.
Anche in questo caso invece di scrivere, ad esempio,
$[(-2,+5)]+[(+2,+4)] = [(-2\cdot+4+2\cdot5,4\cdot5)] = [(2,20)] = [(1,10)]$
scriviamo $-2/5+1/2 = 1/10$.

4. Nal caso dei reali si pone il problema di come si possa rappresentare un numero reale, cioè il problema di trovare un modo per scegliere all'interno di ciascuna classe di successioni di Cauchy equiconvergenti una successione che la rappresenti. Il metodo più utilizzato è quello delle serie di potenze.

Definizione 16.4. Dati a_n, a_{n-1},…,a_0 in $\{0,1,2,3,4,5,6,7,8,9\}$ e $(c_i)_{i\in N}$ in $\{0,1,2,3,4,5,6,7,8,9\}^N$, indichiamo con $a_na_{n-1}...a_0,c_1c_2...$ la successione costituita dai razionali

a_n10^n

$a_n10^n+a_{n-1}10^{n-1}$

. . .

$a_n10^n+a_{n-1}10^{n-1}\ldots+a_010^0$

$a_n10^n+a_{n-1}10^{n-1}\ldots+a_010^0c_110^{-1}$

$a_n10^n+a_{n-1}10^{n-1}\ldots+a_010^0c_110^{-1}\ c_210^{-2}$

. . .

Pertanto $a_na_{n-1}...a_0,c_1c_2...$ denota una serie di potenze di base dieci che a volte per semplicità viene indicata con

$a_n10^n+...+a_010^0+c_110^{-1}+...$

Omettiamo la dimostrazione del seguente fondamentale teorema.

Teorema 16.5. La successione $a_na_{n-1}...a_0,c_1c_2...$ è una successione di Cauchy e quindi individua un numero reale $[a_na_{n-1}...a_0,c_1c_2...]$. Inoltre ogni successione di Cauchy è equiconvergente ad una serie di potenze di base 10.

Definizione 16.6. Diciamo che l'espressione $a_na_{n-1}...a_0,c_1c_2...$ è la *rappresentazione decimale* del numero reale individuato dalla classe $[a_na_{n-1}...a_0,c_1c_2....]$.

Pertanto quando diciamo che il numero π è 3,1415... intendiamo dire che la successione delle somme parziali

$3,$

$3+10^{-1},$

$3+10^{-1}+4\cdot10^{-2}+,$

$3+10^{-1}+4\cdot10^{-2}+10^{-3},$

$3+10^{-1}+4\cdot10^{-2}+10^{-3}+5\cdot10^{-4},$

...

individua una classe di equivalenza che è il numero reale π. Se diciamo che 1/3 è uguale a 0.3333... intendiamo che la successione costantemente uguale ad 1/3 è equiconvergente alla successione 0, $0+3\cdot10^{-1}$, $0+3\cdot10^{-1}+3\cdot10^{-2}$, ...

Da notare che se vogliamo che la rappresentazione sia unica, allora dobbiamo supporre che la prima cifra a_n della parte intera sia diversa da zero. Inoltre dobbiamo evitare le rappresentazioni che usano il periodo nove. Infatti, ad esempio, la successione 2,39999… è equiconvergente alla successione 2,4000…

Problema: Dire quale è il significato dell'asserzione "il numero 0,4399... è uguale al numero 0.44000..." e che cosa si dovrebbe in linea di principio dimostrare per controllare che tale affermazione è vera.

Concludiamo evidenziando che tutto quello che abbiamo detto riferendoci alla base 10 vale anche se si fissa come base un qualunque naturale $b \geq 2$. Pertanto possiamo rappresentare ogni numero reale con una opportuna serie di potenze di base *b*.
Si osservi che poiché $\underline{r}(\underline{r}(x)) = \underline{r}(r([x])) = \underline{r}(x)$.

17. Una costruzione diversa: essere quasi uguali[23]

Una volta che sia stata costruita una estensione del campo ordinato dei razionali che risulta essere un campo ordinato dei numeri reali, abbiamo lo strumento fondamentale con cui costruire buona parte della matematica (tra cui, e non è poco, la geometria analitica e l'analisi matematica). Si pone il problema se sia possibile estendere Q in un diverso campo ordinato per poi eventualmente costruire una matematica diversa. Un possibile tentativo è il seguente.

Il punto di partenza è ancora una volta l'anello $(Q^N, +, \cdot, \underline{0}, \underline{1})$ delle successioni di razionali, ma questa volta:

1. non ci limitiamo alle successioni di Cauchy ma consideriamo la classe di tutte le successioni di razionali
2. introduciamo una relazione di congruenza diversa dalla equiconvergenza.

Precisamente, invece di identificare due successioni che sono "sempre più vicine tra loro", identifichiamo due successioni che sono uguali *"quasi ovunque"* cioè che, in un certo senso, sono uguali tranne che in un *"trascurabile"* insieme di indici. Come dare una formalizzazione tale nozione?

Ricordiamo che in matematica sono spesso utilizzate due nozioni diverse di "quasi ovunque" che sembrano esprimere quest'idea.

[23] Per la comprensione di questo paragrafo è utile andare a vedere l'approccio assiomatico ai numeri reali esposto nel secondo volume.

Esempio 1 di quasi-ovunque. Data una successione $(a_n)_{n\in N}$ viene detto che tale successione verifica quasi-ovunque una proprietà P se $\{n \in N : a_n$ non verifica $P\}$ è finito (e quindi trascurabile). Ad esempio diciamo che $(a_n)_{n\in N}$ è quasi ovunque positiva se $\{n \in N : a_n$ non è positiva$\}$ è finito.

Esempio 2 di quasi ovunque. Sia S un insieme in cui si ha una nozione di misura e sia f una funzione definita in S. Allora diciamo che f verifica quasi ovunque una proprietà P se l'insieme $\{x\in S : f(x)$ non verifica $P\}$ ha misura nulla (quindi trascurabile). Se indichiamo con F la classe degli insiemi *cofiniti* (cioè complementi di finiti), allora nel primo esempio possiamo dire che $(a_n)_{n\in N}$ verifica quasi ovunque P se $\{n\in N : a_n$ verifica $P\} \in F$. Nel secondo esempio possiamo indicare con F la classe degli insiemi l'insieme degli elementi in cui P vale appartiene ad U. Similmente se U denota la classe degli insiemi che sono complementi di insiemi di misura nulla, allora anche nel secondo caso possiamo dire che l'insieme degli elementi in cui vale P appartiene ad U. Inoltre in entrambi i casi la classe U soddisfa le seguenti proprietà.

i) $X\in U$ e $Y\in U \Rightarrow X\cap Y\in U$

ii) $X\in U$ e $Y\supseteq X\Rightarrow Y\in U$.

Tali proprietà suggeriscono la seguente definizione.

Definizione 17.1. Dato un insieme non vuoto S chiamiamo *filtro* una classe U non vuota di sottoinsiemi di S tale che

i) $X\in U$ e $Y\in U \Rightarrow X\cap Y\in U$

ii) $X\in U$ e $Y\supseteq X \Rightarrow Y\in U$.

Diciamo che una proprietà P vale *quasi ovunque* in S rispetto ad U se $\{x\in S :\ x$ soddisfa $P\}\in U$.

Esempio 3. Abbiamo già osservato che la classe dei co-finiti e la classe degli insiemi il cui complemento ha misura nulla sono esempi di filtro. L'intera classe $P(S)$ di sottoinsiemi di S è un filtro che viene chiamato *improprio* (il più grande). Anche il singoletto $\{S\}$ è un filtro (il più piccolo). Fissato un sottoinsieme A di S la classe $U_A = \{X\in \mathcal{P}(S) : X \supseteq A\}$ degli insiemi che contengono A è un filtro che viene detto *principale*. In tale caso P vale quasi ovunque se vale almeno per tutti gli elementi di A. In un certo senso questo vuol dire che si considerano importanti solo gli elementi di A e trascurabili gli altri. Detto in altre parole, “vero quasi ovunque” significa “vero per tutti gli elementi importanti”.

Esercizio. Provare che effettivamente le classi ora definite sono filtri. La classe degli intorni di un punto x nel piano (più in generale in uno spazio topologico) costituisce un filtro. In tale caso una proprietà P vale quasi ovunque se P vale in un intorno di x. Ad esempio due funzioni sono uguali quasi ovunque se coincidono in un intorno di P.

Definizione 17.2. Detto U un filtro nell'insieme N, definiamo in Q^N la relazione binaria $\equiv$ ponendo

$$(a_n)_{n\in N} \equiv (b_n)_{n\in N} \Leftrightarrow a_n = b_n \text{ quasi ovunque rispetto ad } U.$$

Se $(a_n)_{n\in N} \equiv (b_n)_{n\in N}$ diciamo che $(a_n)_{n\in N}$ è *quasi ovunque uguale a* $(b_n)_{n\in N}$.

In altre parole poniamo $(a_n)_{n\in N} \equiv (b_n)_{n\in N}$ se e solo se $\{n\in N : a_n = b_n\}\in U$. Se U è il filtro improprio allora due successioni sono sempre equivalenti. Se U è il filtro $\{N\}$ allora due successioni sono equivalenti solo se coincidono. Se U è il filtro generato dall'insieme $X\subseteq N$, allora due successioni sono equivalenti se almeno in X hanno gli stessi valori.

Teorema 17.3. La relazione $\equiv$ è una congruenza nella struttura $(Q^N, +, \cdot, \underline{0}, \underline{1})$ ed il quoziente $(Q^N/\equiv, +, \cdot, [\underline{0}], [\underline{1}])$ è un anello.

Dim. Proviamo che $\equiv$ è una equivalenza. La proprietà riflessiva e simmetrica sono “ereditate” dalle corrispondenti proprietà dell'uguaglianza. Infatti, poiché $\{n\in N : a_n = a_n\} = N$ ed $N\in U$, risulta che $(a_n)_{n\in N} \equiv (a_n)_{n\in N}$. Assumiamo che $(a_n)_{n\in N} \equiv (b_n)_{n\in N}$ e quindi che $\{n\in N : a_n = b_n\}\in U$. Allora, poiché $\{n\in N : b_n = a_n\} = \{n\in N : a_n = b_n\}$, risulta che $(b_n)_{n\in N} \equiv (a_n)_{n\in N}$. Ciò prova la proprietà simmetrica. Per provare la proprietà transitiva, supponiamo che $(a_n)_{n\in N} \equiv (b_n)_{n\in N}$ e $(b_n)_{n\in N} \equiv (c_n)_{n\in N}$, cioè che $\{n\in N : a_n = b_n\}\in U$ e $\{n\in N : b_n = c_n\}\in U$. In tale caso, poiché

$$\{n\in N : a_n = c_n\} \supseteq \{n\in N : a_n = b_n\}\cap\{n\in N : b_n = c_n\}$$

risulta che $\{n\in N : a_n = c_n\}\in U$ e quindi che $(a_n)_{n\in N} \equiv (c_n)_{n\in N}$.

Per provare che $\equiv$ è una congruenza, supponiamo che $(a_n)_{n\in N} \equiv (a_n')_{n\in N}$ e $(b_n)_{n\in N} \equiv (b_n')_{n\in N}$, cioè che

$\{n\in N: a_n = a'_n\}\in U$ e $\{n\in N: b_n = b'_n\}\in U$.

In tale caso, poiché

$\{n\in N: a_n+b_n = a'_n+ b'_n\} \supseteq \{n\in N: a_n = a'_n\}\cap\{n\in N: b_n = b'_n\}$

possiamo concludere che $(a_n)_{n\in N} + (b_n)_{n\in N} \equiv (a_n')_{n\in N} + (b_n')_{n\in N}$. Esattamente nello stesso modo è possibile provare che $\equiv$ è compatibile con il prodotto.

Per provare che $(Q^N/\equiv , + , \cdot, [\underline{0}], [\underline{1}])$ è un anello osserviamo che, come al solito, il passaggio a quoziente di una struttura conserva tutte le proprietà che si possono esprimere tramite equazioni e che gli assiomi di anello sono espressi tramite equazioni.□

Se $U=\{N\}$ allora $Q^N/\equiv$ coincide con l'anello Q^N. Se U è il filtro principale generato da un singleton allora è evidente che $Q^N/\equiv$ è isomorfo *a* Q.

18. Il campo dei razionali non standard

Se parto da un qualunque filtro non è detto che l'anello $(Q^N/\equiv$, $+ , \cdot , [\underline{0}] , [\underline{1}])$ sia anche un campo U. Basta dire che se il filtro è $\{N\}$ allora $Q^N/\equiv$ è isomorfo a Q^N che abbiamo visto non essere un campo. Le cose non migliorano se si considera il filtro U dei cofiniti. Infatti sia P l'insieme dei numeri pari per cui, ovviamente, $P\notin U$ e $-P\notin U$ e consideriamo la successione $(a_n)_{n\in N}$ tale che $a_n = 1/n$ se $n\in P$ e $a_n = 0$ altrimenti. Allora, poiché $\{n\in N : a_n = 0\} = -P \notin U$, risulta che $[(a_n)_{n\in N}] \neq [\underline{0}]$. Inoltre non esiste nessun elemento $[(b_n)_{n\in N}]$ che sia l'inverso di $[(a_n)_{n\in N}]$. Infatti se fosse $[(a_n)_{n\in N}]\cdot[(b_n)_{n\in N}] = [\underline{1}]$ avremmo che $\{n\in N: a_n\cdot b_n = 1\} \in U$ e, essendo $P \supseteq\{n\in N: a_n\cdot b_n = 1\}$, che che $P\in U$. Potremmo tentare di evitare tale problema considerando un filtro U che contenga P oppure uno che contenga $-P$. Tuttavia l'esempio di un elemento non nullo e non invertibile può essere ripetuto se al posto di P consideriamo un qualunque sottoinsieme X tale che $X\notin U$ e $-X\notin U$. Basta infatti definire $(a_n)_{n\in N}$ponendo

$a_n = 1/n$ se $n \in X$ e $a_n = 0$ altrimenti.

L'unica possibilità è allora quella di considerare filtri tali che, per ogni X, $X\in U$ oppure $-X\in U$.

Definizione 18.1. Diciamo che un filtro proprio U è un *ultrafiltro* se per ogni $X\in P(S)$,

iii) o $X\in U$ oppure $-X\in U$.

Un esempio banale di ultrafiltro si ottiene fissando un elemento $\underline{x}$ di S e ponendo $U_{\underline{x}} = \{X \in P(S) : \underline{x} \in X\}$, considerando cioè il filtro principale generato da un singoletto. La seguente è una caratterizzazione di tali ultrafiltri.

Proposizione 18.2. Sia U un ultrafiltro allora le seguenti asserzioni sono equivalenti:

a) U è principale

b) U è generato da un singleton

c) U contiene un insieme finito.

Oertanto U non è principale se e solo se contiene tutti i cofiniti.

Dim. *a*) $\Rightarrow$ *b*) Sia U un ultrafiltro principale generato da un insieme X e supponiamo che X non sia un singleton. Allora X può essere scomposto in due sottoinsiemi propri e disgiunti X_1 e X_2. Per *iii*) uno di tali due insiemi deve appartenere ad U in contrasto con l'ipotesi che tutti gli elementi di U contengono X.
b) $\Rightarrow$ *c*) Evidente.
c) $\Rightarrow$ *a*) Supponiamo che U sia un ultrafiltro contenente un insieme finito e sia X un insieme che abbia tra tutti i finiti in U il numero minimo di elementi. Allora X non può spezzarsi in due sottoinsiemi propri poiché uno di questi sarebbe un finito più piccolo appartenente ad U. Quindi X è un singleton $\{x\}$. Poiché l'intersezione di due elementi di un filtro proprio è sempre non vuota, un insieme X appartiene ad U se e solo se contiene x.

Infine se l'ultrafiltro U non è principale allora non può contenere nessun insieme finito. Pertanto deve contenere i complementi di tutti gli insiemi finiti, cioè tutti i cofiniti. □

La congruenza associata da un ultrafiltro permette di ottenere un campo.

Teorema 18.3. Sia U un ultrafiltro e $\equiv$ la congruenza generata da U, allora il quoziente di $(Q^N, +, \cdot, \underline{0}, \underline{1})$ modulo $\equiv$ è un campo che estende il campo dei razionali. Denotiamo con Q^* tale campo.

Dim. Per provare che Q^* è un campo supponiamo che $[(a_n)_{n\in N}]$ sia un elemento di Q^* diverso da zero. Allora, poiché lo zero di Q^* è la classe $[(z_n)_{n\in N}]$ determinata dalla successione costantemente uguale a zero, siamo sicuri che $\{n \in N : a_n = 0\} \notin U$. Poiché per ipotesi U è un ultrafiltro, ciò comporta che $\{n \in N: a_n \neq 0\} \in U$. Definiamo la successione $(b_n)_{n \in N}$ ponendo $b_n = 1/a_n$ se $a_n \neq 0$ e $b_n = 1$ altrimenti. Allora, poiché

$$\{n \in N: a_n \cdot b_n = 1\} \supseteq \{n \in N: a_n \neq 0\} \in U,$$

risulta che $\{n \in N: a_n \cdot b_n = 1\} \in U$ e quindi che

$$[(a_n)_{n \in N}] \cdot [(b_n)_{n \in N}] = [(a_n b_n)_{n \in N}] = [\underline{1}].$$

Ciò prova che $[(a_n)_{n \in N}]$ è invertibile.

Per mostrare che Q^* estende Q consideriamo l'applicazione $f: Q \rightarrow Q^*$ che associa ad ogni $r \in Q$ la classe di equivalenza $[\underline{r}]$ dove $\underline{r}$ è la successione costantemente uguale ad r. Tale applicazione risulta essere una immersione di Q in Q^*. □

Se partiamo da un ultrafiltro principale il campo Q^* risulta essere isomorfo al campo dei razionali e quindi non ha alcun interesse.

Proposizione 18.4. Se l'ultrafiltro U è principale allora Q^* è isomorfo a Q.

Dim. Vogliamo provare che l'immersione $f: Q \rightarrow Q^*$ definita nella dimostrazione del teorema precedente è un isomorfismo, cioè è suriettiva. Infatti, sia $\underline{n}$ l'elemento che genera U e sia $[(a_n)_{n \in N}]$ un qualunque elemento di Q^*, allora, se $r = a_{\underline{n}}$, abbiamo che $f(\underline{r}) \equiv (a_n)_{n \in N}$ e quindi $f(\underline{r}) = [\underline{r}] = [(a_n)_{n \in N}]$. □

Come vedremo nel prossimo paragrafo gli ultrafiltri non principali forniscono invece strutture che non sono isomorfe a Q. Si pone tuttavia il problema di dimostrare che esiste effettivamente un ultrafiltro che non è principale. Vale in proposito il seguente teorema di cui omettiamo la dimostrazione in quanto richiede nozioni di teoria degli insiemi che non abbiamo ancora dato.

Teorema 18.5. Se si ammette l'assioma della scelta allora esiste un ultrafiltro che non è principale.

19. Numeri infiniti e numeri infinitesimi in Q^*

Si pone il problema di come si possa definire in Q^* una relazione d'ordine compatibile con le operazioni in Q^*. Poiché Q^* è un'estensione di Q vogliamo trovare una relazione d'ordine che estenda quella che abbiamo definito in Q. Purtroppo mentre sappiamo bene che cosa è una congruenza in una struttura algebrica e come effettuare il quoziente di tale struttura, nel caso siano coinvolte relazioni il discorso è alquanto più problematico.

Fortunatamente, come osservato in Appendice, nel caso di un reticolo possiamo definire la relazione d'ordine facendo riferimento alla sua struttura algebrica. Infatti la struttura ordinata (Q, $\leq$) è un reticolo rispetto alle operazioni $\wedge$ e $\vee$ definite ponendo $x\vee y = max\{x,y\}$ e $x\wedge y = min\{x,y\}$. Tali operazioni si possono estendere a Q^N come abbiamo fatto per l'addizione e la moltiplicazione ponendo

$$(a_n)_{n\in N} \wedge (b_n)_{n\in N} = (a_n \wedge b_n)_{n\in N}$$
$$(a_n)_{n\in N} \vee (b_n)_{n\in N} = (a_n \vee b_n)_{n\in N}.$$

Si ottiene in tale modo un reticolo (Q^N, $\wedge$, $\vee$).

Proposizione 19.1. Dato un ultrafiltro, la relativa relazione di equivalenza $\equiv$ è una congruenza nel reticolo (Q^N, $\wedge$, $\vee$) il cui quoziente ($Q^N/\equiv$, $\wedge$, $\vee$) è ancora un reticolo. Detta $\leq$ la relazione d'ordine associata a tale reticolo, abbiamo che

$$[(a_n)_{n\in N}] \leq [(b_n)_{n\in N}] \Leftrightarrow a_n \leq b_n \text{ quasi ovunque.}$$

Dim. Per provare che $\equiv$ è compatibile con le operazioni $\wedge$ e $\vee$ si procede come fatto nel caso dell'addizione e della moltiplicazione. Che ($Q^N/\equiv$, $\wedge$, $\vee$) sia un reticolo segue dal fatto che gli assiomi della teoria dei reticoli sono equazioni e che le equazioni si conservano quando effettuiamo una potenza diretta ed un quoziente. Infine abbiamo che

$$\begin{aligned}[(a_n)_{n\in N}] \leq [(b_n)_{n\in N}] &\Leftrightarrow [(a_n)_{n\in N}] \wedge [(b_n)_{n\in N}] = [(a_n)_{n\in N}].\\ &\Leftrightarrow min\{a_n,b_n\} = a_n \text{ quasi ovunque}\\ &\Leftrightarrow a_n \leq b_n \text{ quasi ovunque.}\end{aligned}$$ □

Definizione 19.2. Sia U un ultrafiltro, allora indichiamo con Q^* la struttura ($Q^N/\equiv$, $+$, $\cdot$, $[\underline{0}]$, $[\underline{1}]$, $\leq$) e diciamo che Q^* è un *campo di numeri razionali nonstandard.*

Un aspetto particolarmente interessante dei campi di tale tipo è quello di essere non archimedeo e quindi di contenere infiniti ed infinitesimi (si veda in Appendice).

Teorema 19.3. Se l'ultrafiltro U non è principale allora il campo dei numeri razionali non standard Q^* non è archimedeo.

Dim. Consideriamo una qualunque successione positiva divergente, ad esempio la successione $(n^2)_{n\in N}$ ed il corrispondente numero iperreale $[(n^2)_{n\in N}]$. Allora preso un qualunque intero p risulta che $X = \{n\in N : n^2 < p\}$ è finito e quindi che $\{n\in N : n^2 \geq p\}$ è cofinito. Poiché abbiamo supposto che U contiene tutti i cofiniti, tale insieme appartiene ad U. Questo prova che $[(n^2)_{n\in N}] \geq p\cdot[\underline{1}]$ qualunque sia p, cioè che non esiste nessun multiplo di $[\underline{1}]$ capace di superare $[(n^2)_{n\in N}]$. □

Ovviamente l'nverso $[(1/n^2)_{n\in N}]$ di $[(n^2)_{n\in N}]$ è un infinitesimo.

Corollario 19.4. Se l'ultrafiltro U non è principale allora il campo ordinato Q^* non è isomorfo al campo ordinato Q.

Dim. Basta osservare che strutture isomorfe verificano le stesse proprietà e che mentre Q è archimedeo Q^* non lo è. □

Chiudiamo questo paragrafo esaminando rapidamente la questione della radice quadrata di 2 cioè il problema di trovare un numero x tale che $x^2 = 2$. Ebbene mentre tale problema è risolvibile nel campo dei reali, non lo è invece in Q^*.

Proposizione 19.5. In Q^* l'equazione $x^2 = 2$ non ammette soluzioni, in altre parole in Q^* non esiste la radice di 2. Tuttavia se U non è principale allora esiste $x\in Q^*$ tale che x^2 è *infinitamente vicino* a 2, cioè tale che la differenza x^2 -2 è un infinitesimo.

Dim. Supponiamo che esista $[(a_n)_{n\in N}]$ in Q^* tale che $[(a_n)_{n\in N}]\cdot[(a_n)_{n\in N}] = [2]$. Allora dovrebbe essere che $a_n^2 = 2$ quasi ovunque cosa impossibile poiché nessun numero razionale verifica tale uguaglianza (la radice quadrata di 2 non è razionale).

Per trovare un iperrazionale x tale che x^2 è infinitamente vicino a 2, consideriamo una successione $(q_n)_{n\in N}$ di razionali tali che $lim_{n\to\infty}\, q_n^2 = 2$.[24] Allora per ogni naturale non nullo p risulta

[24] L'esistenza di una tale successione è ben nota agli studenti delle superiori che hanno imparato ad estrarre le radici quadrate e quindi, in particolare, a calcolare la radice di due. Tale calcolo consiste infatti nella produzione di una successione $(q_n)_{n\in N}$ di razionali, in forma decimale, il cui quadrato è sempre più vicino a 2.

che $\{n \in N : |q_n^2 - 2| \leq 1/p \cdot 1\}$ è cofinito e quindi appartiene ad U. Questo prova che $[(q_n)_{n\in N}]^2$ è infinitamente vicino a 2. □

Il campo dei reali non standard. Non andiamo oltre a tali questioni perché in questo libro lo scopo di introdurre Q^* è solo quello di fare intravedere un universo nuovo ed interessante che è quello dell'*analisi non standard*. Ci limitiamo ad osservare che con la stessa tecnica con cui abbiamo costruito il campo Q^* dei razionali non standard a partire dal campo dei razionali possiamo costruire il campo R^* dei reali non standard a partire dal campo dei reali. Basta considerare successioni di numeri reali invece che successioni di razionali.

Teorema 19.6. Sia U un ultrafiltro che contiene il filtro dei cofiniti e definiamo in R^N la relazione $\equiv$ ponendo

$$(a_n)_{n\in N} \equiv (b_n)_{n\in N} \Leftrightarrow \{n \in N : a_n = b_n\} \in U.$$

Allora $\equiv$ è una congruenza nella struttura $(R^N, +, \cdot, \underline{0}, \underline{1})$. Il relativo quoziente $(R^N/\equiv, +, \cdot, [\underline{0}], [\underline{1}])$ è un campo non archimedeo rispetto all'ordinamento definito come in proposizione 19.1.

Tale campo prende il nome di *campo dei numeri reali non standard*. L'approccio all'analisi matematica che utilizza un campo di numeri reali non standard prende il nome di *Analisi non standard.* In un certo senso l'analisi non standard è un modo puramente algebrico di trattare il calcolo differenziale. Infatti la presenza di infiniti ed infinitesimi permette, per fare un esempio, di definire il limite di una funzione $f(x)$ per x che tende all'infinito direttamente come il valore di f calcolato in un numero infinito. Un integrale definito è effettivamente una sorta di somma infinitaria, la derivata è il rapporto tra due infinitesimi e così via. Ciò spesso fornisce metodi eleganti e naturali per dimostrare teoremi in analisi matematica.

20. Invertire il percorso: dal continuo geometrico ai numeri

In questo capitolo abbiamo visto che buona parte della matematica nasce dalla seguente catena di costruzioni successive:

naturali → relativi → razionali → reali → geometria →...

Questa catena basa tutto sulla nozione di numero naturale che si basa sull'attività del contare, uno dopo l'altro, gli elementi di un insieme finito di oggetti. Questa è un'idea profondamente radicata in ciascun uomo che, come abbiamo visto, si può formalizzare nella teoria delle terne di Peano. D'altra parte anche l'idea di continuo geometrico è un'idea profondamente radicata nell'uomo e quindi si pone il problema se sia possibile invertire la catena di costruzioni che è stata oggetto di questo capitolo al modo seguente:

geometria $\rightarrow$ *reali* $\rightarrow$ *razionali* $\rightarrow$ *relativi* $\rightarrow$*naturali*.

Questo modo di procedere in effetti lo abbiamo già trovato in modo implicito nella teoria delle grandezze omogenee di Eudosso-Euclide in cui la nozione di proporzione contiene in se le operazioni di prodotto e divisione. Infatti, scelta una grandezza u come unitaria, la proporzione $a:b = u:c$ ci dice che $b = b\cdot u = a\cdot c = b$. La proporzione $b : u = a:c$ ci dice che $b = a:c$. Naturalmente perché il tutto funzioni ci si deve assicurare

- l'esistenza e l'unicità di una grandezza x tale che $a:x = u:c$.
- l'esistenza e l'unicita di una grandezza x tale $x:u = a:c$.

Questo viene espresso col dire che fissate tre grandezze esiste ed è unica una grandezza quarta proporzionale.

In modo più esplicito lo abbiamo trovato nel calcolo dei segmenti di Cartesio agli inizi della geometria analitica. Da tenere presente che in entrambi i casi si considerano tutte grandezze positive quindi siamo lontani dall'ottenere un campo.

Tuttavia proprio Gottlob Frege, uno dei creatori della catena che va dall'aritmetica alla geometria, negli ultimi anni della sua vita capovolge completamente i suoi convincimenti asserendo decisamente la necessità di dare un fondamento geometrico all'idea di numero. Egli infatti dopo aver conosciuto il paradosso di Russell che mostrava la contradditorietà della teoria degli insiemi affermerà:

Nulla di più indesiderabile può capitare ad un ricercatore del fatto che una delle fondamenta del suo edificio si incrini dopo che l'opera è finita. E questa la situazione in cui mi trovo in seguito ad una lettera (contenente il paradosso) inviatami del sig. Bertrand Russell proprio mentre si stava ultimando la stampa di questo volume. (*Fondamenti dell'aritmetica"* nel 1903)

Passeranno un po' di anni e in uno scritto degli anni 1924-25 Frege dichiarerà:

Più rifletto su questo punto più mi convinco che l'aritmetica e la geometria sono cresciuti dallo stesso terreno, precisamente quello della geometria, in modo che tutta la teoria dei numeri è, propriamente, geometria. La matematica appare quindi perfettamente unitaria nella sua essenza. Il contare, che ha avuto origine dalle esigenze della vita pratica, ha ingannato gli studiosi.

Purtroppo Frege nel 1925 muore prima di poter sviluppare questo suo nuovo punto di vista. Tuttavia possiamo ritrovare una costruzione di natura geometrica dei numeri reali, delle operazioni aritmetiche della relazione d'ordine nei famosi *Fondamenti* della geometria di Hilbert.

Non mi soffermerò ulteriormente sulla possibilità di definire i numeri reali in modo geometrico pur essendo questo un argomento che mi sembra di grande interesse.[25]

[25] Suggerisco comunque di leggere l'articolo di G. Anatriello, su una Fondazione Geometrica della Matematica, *Science & Philosophy*, Vol. 5(1), 2017, pp. 91-108.

LETTURA

Cesare Zavattini, *Gara di Matematica*: da *I tre libri, Parliamo tanto di me* - Bompiani - cap. *XVI* pag. 48,49,50.[1]

E' un ricordo della mia infanzia. Abitavo a Gottinga nel dicembre del milleottocentosettanta. Mio padre ed io giungemmo all'Accademia quando il presidente Maust stava cominciando l'appello dei partecipanti alla *Gara Mondiale di Matematica*. Subito babbo andò a mettersi fra gli iscritti dopo avermi affidato alla signora Katten, amica di famiglia. Seppi da lei che il colpo del cannone di Pombo, il bidello, avrebbe segnato l'inizio della storica contesa. La signora Katten mi raccontò un episodio, ignoto ai più, intorno all'attività di Pombo. Costui sparava da trent'anni un colpo di cannone per annunciare il mezzogiorno preciso. Una volta se n'era dimenticato. Il dì appresso, allora, aveva sparato il colpo del giorno prima, e così di seguito fino a quel venerdì del milleottocentosettanta. Nessuno a Gottinga si era mai accorto che Pombo sparava il colpo del giorno avanti.[2]

Esauriti i preliminari, la gara ebbe inizio alla presenza del principe Ottone e di un ragguardevole gruppo di intellettuali.

"Uno, due, tre, quattro, cinque …"

Nella sala si udivano soltanto le voci dei gareggianti. Alle diciassette circa, avevano superato il ventesimo migliaio. Il pubblico si appassionava alla nobile contesa e i commenti si intrecciavano. Alle diciannove, Alain, della Sorbona, si accasciò sfinito.

Alle venti, i superstiti erano sette.

"36767, 36768, 36769, 36770…"

Alle ventuno Pombo accese i lampioni. Gli spettatori ne approfittarono per mangiare le provviste portate da casa.

"40719, 40720, 40721…"

Io guardavo mio padre, madido di sudore, ma tenace. La signora Katten accarezzandomi i capelli ripeteva come un ritornello: 'Che bravo babbo hai,' e a me non pareva neppure di avere fame. Alle ventidue precise avvenne il primo colpo di scena: l'algebrista Pull scattò:

---

[1] Esponiamo un breve racconto di Zavattini che in un certo senso parla dell'infinito e dell'operatore di successore (e quindi della natura di una terna di Peano).

[2] Tale episodio può essere visto come il fatto che per un cardinale infinito x accade che $x+1 = x$ (si veda il volume 2 nel capitolo che riguarda la cardinalità).

"Un miliardo"
Un *oh* di meraviglia coronò l'inattesa sortita; si restò tutti col fiato sospeso. Binacchi, un italiano, aggiunse issofatto:
"Un miliardo di miliardi di miliardi."
Nella sala scoppiò un applauso subito represso dal Presidente. Mio padre guardò intorno con superiorità, sorrise alla signora Katten e cominciò:
"Un miliardo di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi…"
La folla delirava:
"Evviva, evviva."
La signora Katten e io, stretti uno all'altro, piangevamo dall'emozione.
"…di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi."
Il presidente Maust, pallidissimo, mormorava a mio padre, tirandolo per le falde della palandrana:
"Basta, basta, le farà male."
Mio padre seguitava fieramente:
"… di miliardi di miliardi di miliardi di miliardi ..."
A poco a poco la sua voce si smorzò, l'ultimo fievole *di miliardi* gli uscì dalle labbra come un sospiro, indi si abbattè sfinito sulla sedia. Gli spettatori in piedi lo acclamavano freneticamente. Il principe Ottone gli si avvicinò e stava per appuntargli una medaglia sul petto quando Gianni Binacchi urlò:
"Più uno!"
La folla precipitatasi nell'emiciclo portò in trionfo Gianni Binacchi. Quando tornammo a casa, mia madre ci aspettava ansiosa alla porta. Pioveva. Il babbo, appena sceso dalla diligenza, le si gettò tra le braccia singhiozzando:
"Se avessi detto più due avrei vinto io".[3]

[3] Il babbo avrebbe anche potuto esclamare "il successivo dell'ultimo numero detto da Binacchi!" e sedersi tranquillo ad aspettare. Si sarebbe creata una situazione drammatica per Binacchi. Per quanto fosse caparbio nel contare prima o poi sarebbe stato costretto a fermarsi. Ed allora il babbo avrebbe vinto! Le cose sarebbero diventate un po' più complicate se Binacchi nel momento di fermarsi avesse esclamato "il successivo del numero indicato dal babbo!". A questo punto sarebbe stato davvero difficile per il Presidente decidere chi fosse il vincitore.

CAPITOLO 4

DEFINIRE LA NOZIONE DI ALGORITMO

> *"Non è infatti degno di uomini di ingegno perdere ore come schiavi nel lavoro di calcolo che potrebbe essere affidato tranquillamente a chiunque altro se si usassero le macchine"* (Gottfried Leibnitz, seconda metà del '600).

1. Un po' di storia: gli abachi

Per l'uomo comune spesso "essere bravo in matematica" è sinonimo di "sapere fare bene i calcoli". Come invece è noto a tutti quelli che hanno bazzicato una facoltà scientifica o un buon liceo, "essere bravo in matematica" significa piuttosto "essere capace di dimostrare teoremi" ed "essere capace di trovare metodi generali per risolvere problemi". Questa visione della matematica l'abbiamo ereditata dalla cultura aristocratica dei antichi greci. D'altra parte, paradossalmente, tale visione si è dimostrata potentissima proprio nelle applicazioni e non può che essere condivisa. Tuttavia la dimensione "calcolo" quando si parla di matematica non può essere completamente esclusa, specialmente se si tiene conto dell'attuale sviluppo del ruolo dei calcolatori.

E' probabile che le dita delle mani siano state i primi strumenti materiali per rappresentare i numeri e per facilitare l'esecuzione dei calcoli numerici. Non è un caso che la numerazione attuale sia in base dieci. Anche il numero 12 gioca un ruolo (spesso le uova si sono vendute e si vendono a dozzine) poiché è possibile contare su una mano fino a dodici poggiando il pollice sulle tre parti in cui sono suddivise le rimanenti dita. Successivamente, per rappresentare numeri per i quali le dita o altre parti del corpo non possono essere sufficienti è stato necessario usare altri oggetti come tacche su ossa o pezzi di legno, sassolini, conchiglie, ed altro[4].

Ben presto tali modi di rappresentare i numeri diventarono

[4] Di questa origine del calcolo rimane traccia nelle parole usate in informatica e matematica. Infatti:
- dal nome latino delle dita, *digita*, trae origine l'aggettivo *digitale*,
- dal nome latino dei piccoli sassi, *calculi*, deriva il termine *calcolo*.

anche uno strumento per eseguire le operazioni elementari di addizione, sottrazione e moltiplicazione. Apparve così *l'abaco* che probabilmente fu il primo strumento costruito dall'uomo per eseguire calcoli. Gli abachi venivano costruiti con i materiali più diversi e se ne trovano tracce fin dal 3000 A. C. in Mesopotamia. Per dare un idea del suo funzionamento, descriviamo l'abaco russo che ha la stessa struttura del pallottoliere che a volte si regala ai bambini. Esso è costituito da una serie di aste metalliche parallele in ciascuna delle quali sono infilate dieci palline. Le linee da destra verso sinistra rappresentano le unità, le decine, le centinaia e così via.

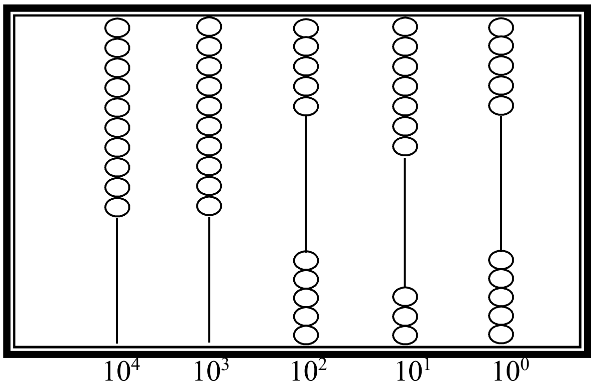
Nell'abaco un numero intero $c_n...c_3c_2c_1c_0$, scritto in base dieci è rappresentato dallo spostare verso il basso:

- c_0 palline nella linea delle unità
- c_1 palline nella linea delle decine
- c_2 palline nella linea delle centinaia

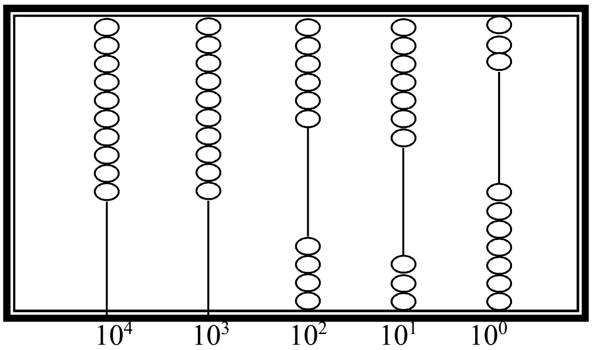
. . .

Abaco Russo

Ad esempio le seguenti due configurazioni rappresentano il numero 535 ed il numero 437, rispettivamente.



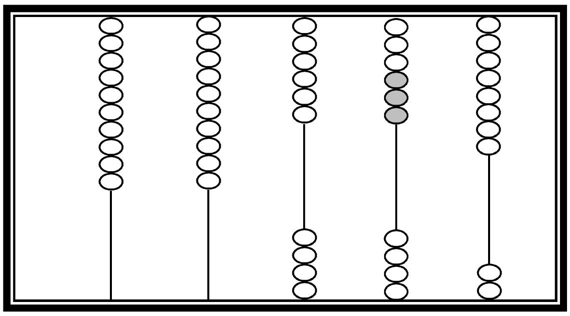
535



437

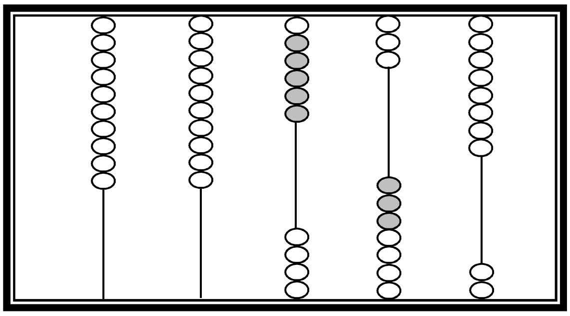
Una tale rappresentazione suggerisce poi un modo per effettuare l'addizione. Partiamo ad esempio da 437. Se dovessi aggiungere 2 non ci sarebbero problemi: dovrei abbassare due palline della colonna delle unità. Tuttavia se dovessi aggiungere 6 avendo a disposizione solo 3 palline da abbassare avrei delle difficoltà. Tuttavia aggiungere 6 significa aggiungere 10 e togliere 4, e quindi aggiungere in basso un palllina delle decine e togliere 4 palline dalla colonna delle unità. Dovendo ad esempio addizionare a 437 il numero 535, si procede al modo seguente:

1. Per addizionare le 5 unità di 535 si abbassa una pallina delle decine e si tolgono 10-5 = 5 palline dell'unità. Si arriva alla seguente configurazione:



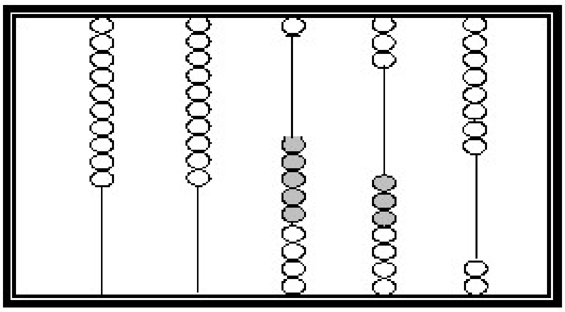
437+5

2. Per addizionare le 3 decine di 535 si opera sulla colonna delle decine spostando verso il basso 3 palline



437+35

3. Infine si passa a lavorare nella colonna delle centinaia spostando verso il basso 5 palline



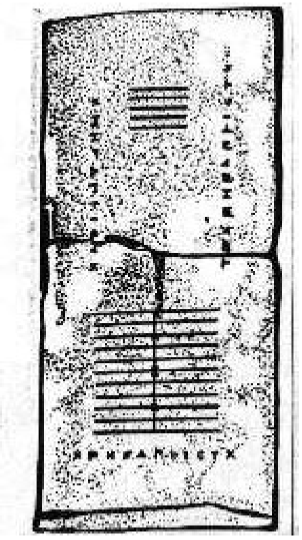
437+535

La cosa, più difficile da descrivere che da fare, si traduce in una sequenza di azioni che si eseguono meccanicamente e rapidamente.

Da notare che non ha importanza il fatto che materialmente la procedura è eseguita dall'uomo. Siamo in presenza di una vera e propria macchina calcolatrice perché la sequenza di azioni da compiere potrebbe essere eseguita anche da una persona che non conosce il significato dei numeri.

2. Vari tipi di abaco

Strumenti simili agli abachi sono presenti in tutte le culture ed in tutti i tempi. Da notare che gli abachi erano utilizzati molto prima che fosse inventata la rappresentazione posizionale dei numeri e la stessa numerazione scritta. Inoltre, nonostante la loro origine antichissima continuano ad essere utilizzati sino ai tempi attuali. Questo sia perché un calcolo che si sviluppa tramite movimenti è più rapido di quello "con carta e penna", sia perché tale tipo di calcolo non comporta il saper leggere e scrivere[5].

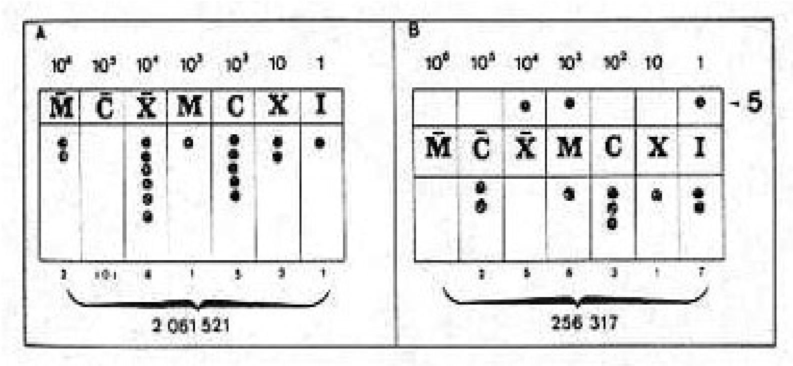


Abaco Babilonese
Del V, IV secolo a.C.

Nella figura affianco è riprodotto un esemplare di abaco dei Babilonesi, i quali intorno al V, IV sec. a.C. utilizzavano già uno strumento di marmo di forma rettangolare, su cui erano incisi due gruppi di undici linee verticali attraversate da una linea orizzontale.

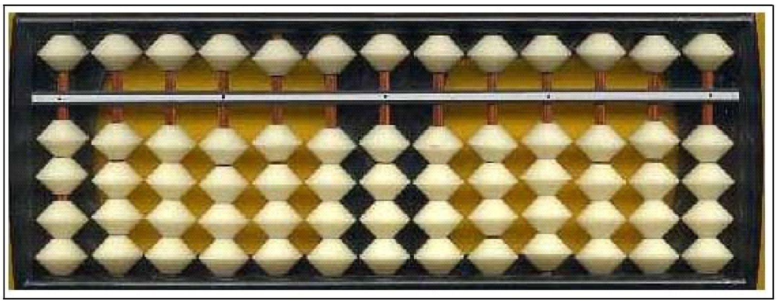
Sotto sono riprodotti due abachi di epoca romana. Il primo non è diverso dal pallottoliere che abbiamo già visto. Infatti nel disegno il numero 2051521 è rappresentato come al solito come somma di 1 unità, 2 decine e così via. Il secondo invece è diviso in una striscia inferiore in cui venivano posti al massimo quattro sassolini ed una striscia superiore in cui veniva posto al più un sassolino. I sassolini posti nella sezione inferiore rappresentano le unità, le decine, e le successive potenze di dieci. Quelli della sezione superiore valevano invece cinque volte quelli delle sezioni inferiori della colonna corrispondente. In tale tipo di abaco

[5]Si deve tenere presente che fino a tempi abbastanza recenti erano poche le persone che sapevano leggere, e che le cose non è che siano poi tanto migliorate nei tempi attuali se si va al di fuori dei paesi occidentali.



Abachi Romani

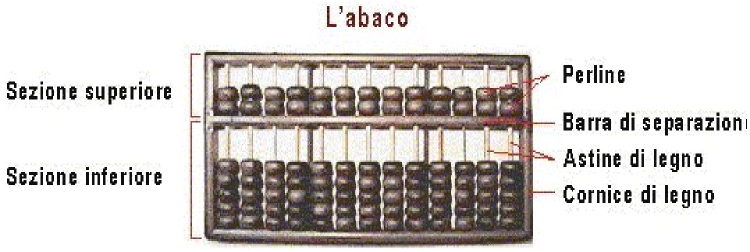
il numero 256317 è rappresentato come
- [2 unità ed 1 di cinque] +
- [1 di dieci] +
- [3 di cento] +
- [1 di mille ed 1 di cinquemila] +
- [1 di cinquantamila] +
- [2 di centomila]



Abaco giapponese

La stessa idea è alla base degli abachi giapponesi, ancora in uso, che utilizzano palline infilate in aste e non sassolini poggiati su pietra o argilla.
Leggermente diverso è l'abaco cinese che riportiamo qui sotto. Esso presenta in ogni linea verticale 5+2 invece che 4+1 palline. Come abbiamo visto, per rappresentare i numeri sarebbe bastata una sola perlina nella parte superiore, e quattro nella parte inferiore. Le due perline in più servono come memoria aggiuntiva

per operazioni più complesse.



Abaco cinese

Gli abachi sono ancora usati nei mercati di molte popolazione arabe e fino a pochi anni fa erano ancora usati in Russia. Anche i cinesi e i giapponesi erano molto abili nel calcolo strumentale e ancor oggi la cultura dell'abaco è molto diffusa in Cina e in Giappone[6].

3. I bastoncini di Nepero.

Uno dei forti limiti degli abachi è la difficoltà di effettuare le moltiplicazioni le quali dovevano essere eseguite, in modo alquanto complicato, come addizioni iterate. Macchine per effettuare le moltiplicazioni sono legate ad un metodo antico di calcolo diffuso presso gli arabi che viene detto a *'graticola'* o a *'gelosia'*. Tale denominazione viene spiegata da Luca Pacioli nella S*umma de Aritmetica* del 1494 al modo seguente:

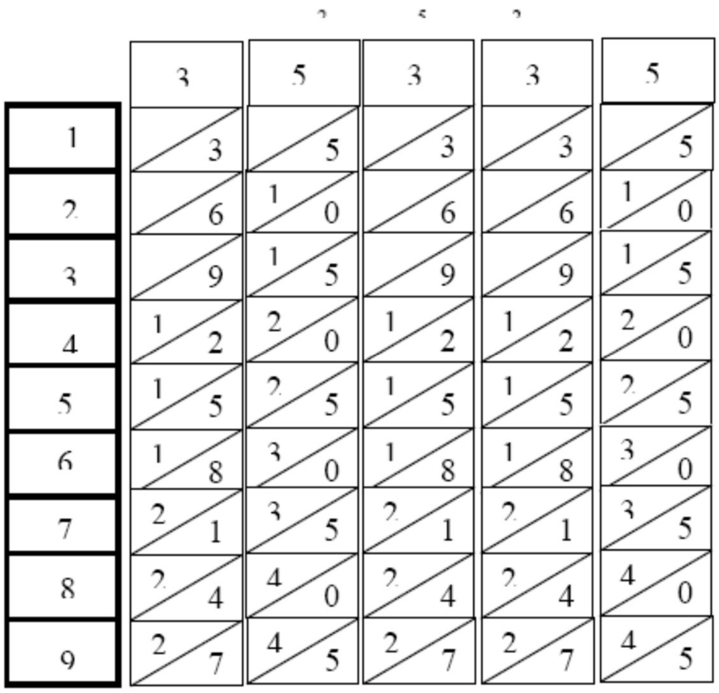
' ...per gelosia intendiamo quelle graticelle che si costumano mettere alle finestre de le case dove habitano done, acio che non si possino facilmente vedere.'

Ecco il metodo: dovendo moltiplicare due numeri come 252 e 347, si costruisce un rettangolo diviso in nove rettangoli più piccoli. Ciascun rettangolo viene diviso in due parti da una diagonale come indicato in figura. I numeri da moltiplicare si pongono sui due lati del rettangolo.

1

6 Il 12 novembre 1946, cioè agli albori dei calcolatori elettronici, a Tokyo si tenne una curiosa competizione sportiva di velocità fra un operatore di calcolatrice elettrica e un abachista, con un risultato davvero incredibile: la palma della vittoria spettò all'abachista!

All'interno dei rettangoli piccoli viene posto il prodotto dei due numeri corrispondenti, scrivendo la parte intera al di sotto della diagonale e le decine al di sopra. Il risultato si ottiene sommando i contenuti delle strisce trasversali individuate dalle diagonali (strisce che corrispondono alle unità, alle decine, alle centinaia ...) e tenendo conto dell'eventuale riporto.



Tale metodo fu ripreso dallo scozzese Giovanni Nepero (1550-1617) famoso per avere introdotto l'uso dei logaritmi e delle relative tavole di esecuzione dei calcoli. Nepero propose però di utilizzare al posto di un disegno su di un foglio una sequenza di bastoncini. Ciascun bastoncino corrispondeva alla moltiplicazione di un numero di una sola cifra per un qualunque numero da 0 a nove. Pertanto esistevano dieci tipi di bastoncini, ciascuno diviso in 10 quadrati. Questi quadrati, a loro volta, erano tagliati da una diagonale, sopra la quale stavano i numeri delle decine, mentre sotto stavano i numeri delle unità.

Facciamo un esempio: moltiplichiamo 7×35335. Accanto al regolo fisso si pongono i regoli 3, 5, 3, 3, 5. Poi si legge nella riga 7 del regolo fisso. Si leggerà da destra verso sinistra, 5 unità + (1+3) decine (1+2) centinaia +(5+2) migliaia +(1+3) di diecimila

+ 2 di centomila ed il risultato sarà 247345. Può presentarsi anche il problema del riporto. Ad esempio se si vuole 6×35335 allora, leggendo da destra verso sinistra, si avranno

0	unità +
8+3	decine +
8+1	centinaia +
0+1	migliaia +
8+3	di diecimila +
1	di centomila

Il tutto è uguale a 212010.

Naturalmente non è difficile leggere il risultato direttamente da sinistra verso destra. Ad esempio 5 × 35335 si può leggere direttamente come il numero 176675. In un certo senso siamo in presenza di una sorta di estensione della tavola pitagorica.

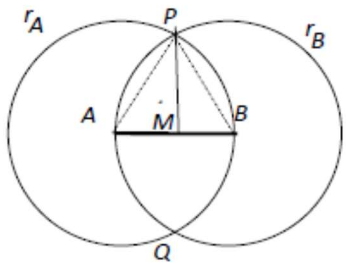
Sopra è riportata una foto di una antica scatola di bastoncini di Nepero. Ecco invece una moderna realizzazione in legno.

4. Un antico calcolatore analogico: riga+compasso
Gli abachi ed i bastoncini di Nepero potevano servire per il calcolo delle grandezze intere oppure, utilizzando la virgola, per i numeri decimali finiti. Per il calcolo delle grandezze continue l'unica cosa era invece rivolgersi alle risorse della geometria ed, in particolare, ai due strumenti fondamentali dagli antichi greci: la riga ed il compasso. Quando abbiamo parlato del calcolo dei segmenti proposto da Cartesio abbiamo già visto come sia possibile effettuare le quattro operazioni aritmetiche e la radice quadrata sui segmenti. Più complessa la questione se vogliamo effettuare calcoli su altre grandezze geometriche, ad esempio calcoli su angoli, aree o volumi.

Esaminiamo ad esempio i calcoli da fare sugli angoli. Le operazioni di addizione e sottrazione di due angoli non presentano grosse difficoltà poiché si basano, come nel caso dei segmenti, sulla possibilità di spostare due angoli senza che la loro grandezza sia alterata. Per lo stesso motivo non è difficile raddoppiare o triplicare un triangolo. Non altrettanto facile è la divisione in parti uguali. Cominciamo con il problema della divisione per due.

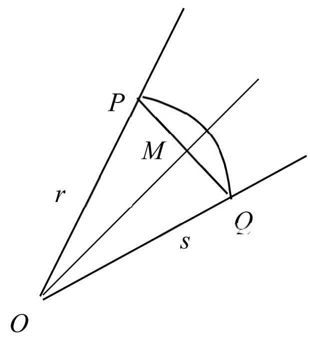
Proposizione 4.1. I segmenti e gli angoli possono essere divisi in due parti uguali usando la riga ed il compasso[7].

Dim. Naturalmente per bisecare un segmento potremmo utilizzare il metodo di Cartesio basato sulle proporzioni esposto nel capitolo 2. Un altro modo è di utilizzare la costruzione di un triangolo equilatero di un dato lato che abbiamo già visto quando abbiamo parlato di Euclide. Infatti dato un segmento *AB* possiamo centrare il compasso in *A*, aprirlo



fino a raggiungere *B* e poi tracciare la circonferenza r_A. Successivamente centrare in *B*, aprire fino a raggiungere *A* e tracciare la circonferenza r_B. Detto *P* uno dei due punti di intersezione delle circonferenze, il punto medio *M* si ottiene intersecando il segmento *AB* con la perpendicolare dal punto *P*. Infatti i due triangoli rettangoli *APM* e *BPM* sono uguali essendo rettangoli ed essendo *AP* = *PB*.

Nel caso della bisezione dell'angolo è sufficiente ricondurre il problema a quello della bisezione di un segmento. Infatti, supponiamo che α sia un angolo racchiuso tra le rette *r* ed *s* che si incontrano nel punto *O*. Possiamo allora tracciare una qualunque circonferenza di centro *O*. Detti *P* e



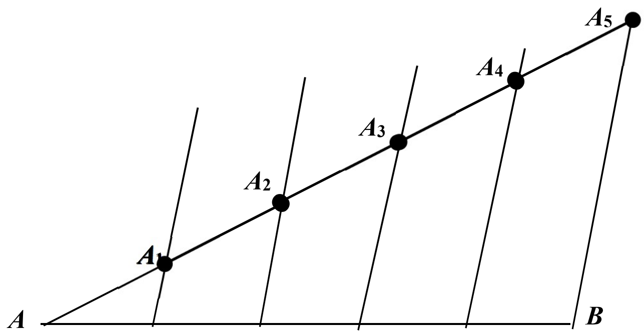
Q i punti di intersezione di tale circonferenza con le due rette e calcolato (per meglio dire costruito) il punto medio *M* di tale segmento, la retta *OM* divide in due parti l'angolo α. □

La bisezione di un segmento si ottiene anche dalla la seguente proposizione al caso $n = 2$.

[7] Nella teoria delle costruzioni con riga e compasso si precisa che la riga non è graduata e che il compasso non può essere usato per riportare distanze da una parte all'altra del piano ma solo per tracciare circonferenze.

Proposizione 4.2. Dato un segmento *AB* ed un numero naturale *n* è possibile costruire con riga e compasso un segmento di lunghezza uguale alla *n*-sima parte di *AB*.

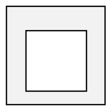
Dim. Tracciamo una retta per *A* diversa da *AB* e, supponendo ad esempio $n = 5$, fissiamo i punti A_1, A_2, A_3, A_4, A_5 in modo che $AA_1 = A_1A_2 =$, $A_2A_3 = A_3A_4 = A_4A_5$ (è sufficiente un compasso). Successivamente tracciamo quattro rette parallele ad A_5B per i punti A_1, A_2, A_3, A_4, rispettivamente. E' evidente che le intersezioni di tali rette con *AB* forniscono segmenti la cui lunghezza è la quinta parte di *AB*.



□

Purtroppo non possiamo utilizzare tale risultato per la divisione di un angolo in n parti uguali come abbiamo fatto per la bisezione. Infatti, a differenza del caso $n = 2$, se anche dividiamo il segmento PQ in tre parti uguali tracciando i punti M_1 e M_2, le rette OM_1 e OM_2 non dividerebbero l'angolo in tre parti uguali.

Problema: Costruire un quadrato equivalente alla parte grigia della figura affianco (si utilizzi il teorema di Pitagora).



Problema: E' possibile dividere un angolo per 8 ?

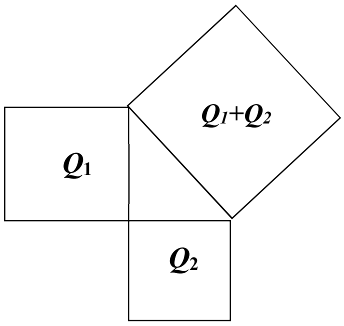
5. Teoremi limitativi per la macchina riga+compasso: altre macchine

Il problema di dividere un angolo in n parti uguali rientra in quelli che spesso vengono chiamati *"i tre problemi famosi dell'antichità"* che sono :

- La trisezione dell'angolo
- La duplicazione del cubo[8]
- La rettificazione della circonferenza.

[8] Circa l'origine del problema della duplicazione del cubo si racconta di una terribile epidemia scoppiata ad Atene che spinse i suoi abitanti a chiedere aiuto all'oracolo di Delfi. La risposta dell'oracolo fu "Se volete placare le ire divine raddoppiate l'altare di Apollo ". Tale altare aveva forma cubica e ciò, appunto, dette origine al problema.

Abbiamo già visto che la bisezione di un angolo con riga e compasso è abbastanza semplice. Anche la duplicazione del quadrato non presenta molte difficoltà.



Proposizione 5.1. Dati due quadrati Q_1 e Q_2 è sempre possibile costruire un quadrato la cui area sia la somma delle aree di Q_1 e Q_2. Pertanto il problema della duplicazione del quadrato è risolvibile con riga e compasso.

Dim. Per il teorema di Pitagora basta disporre i due quadrati in modo che i relativi lati siano i cateti di un triangolo rettangolo. Il quadrato costruito sull'ipotenusa è quello cercato.

Per la duplicazione del quadrato basta considerare due quadrati uguali. □

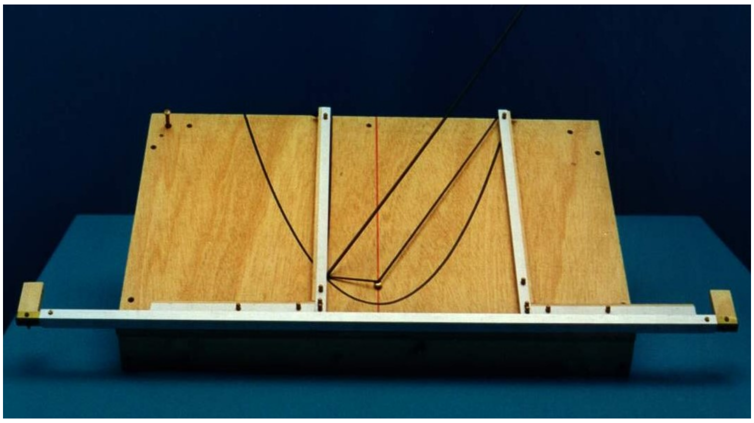
Purtroppo per i tre problemi che abbiamo elencato le cose non sono altrettanto semplici. Vale infatti il seguente teorema di cui omettiamo la dimostrazione che richiederebbe nozioni relative agli ampliamenti dei campi.

Teorema 5.2. (Teoremi limitativi per la riga ed il compasso). Nessuna costruzione con riga e compasso permette di costruire un angolo che sia la terza parte di un dato angolo α[9]. La stessa cosa può essere affermata per il problema della duplicazione del cubo e per quello della rettificazione della circonferenza.

Sembra pertanto che la macchina riga+compasso non sia abbastanza potente.

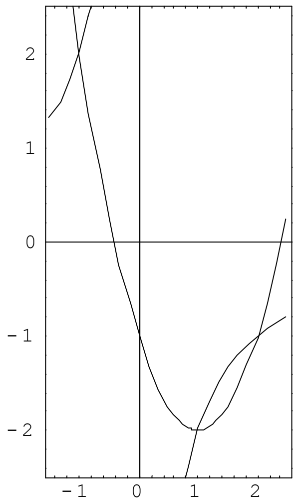
---

[9] E' necessario distinguere il problema dell'esistenza di un segmento dal problema della sua costruzione con riga e compasso. Entrambi i problemi non sono semplici. Ad esempio, data una circonferenza non si può dare per scontato che esista un segmento della stessa lunghezza. Tuttavia se si accetta l'assioma di continuità (di cui parleremo in seguito) non è difficile dimostrare che tale segmento esiste. D'altra parte è stato dimostrato che esso non può essere costruito con riga e compasso. Discorso analogo vale per gli altri problemi dell'antichità.



Possiamo allora tentare di migliorare la situazione utilizzando altre macchine che tracciano curve, ad esempio coniche. Tali macchine esistevano. Ad esempio, nella foto è riportata una macchina per tracciare parabole[10].

Non è difficile vedere che tali macchine permettono di costruire le radici delle equazioni di secondo, terzo e quarto grado. L'idea è la seguente: usualmente, dovendo trovare i punti di intersezione di due curve di cui si conosce l'espressione analitica, applichiamo il seguente procedimento:



- scriviamo il sistema delle relative equazioni (in due variabili e di secondo grado),
- calcoliamo l'equazione risultante eliminando una variabile (ottenendo una equazione in una sola variabile di quarto grado),
- risolviamo tale equazione con qualche formula.

Le soluzioni ottenute saranno le ascisse dei punti d'intersezione. Noi invece invertiamo tale procedimento partendo da una equazione di quarto grado in una va-

[10] Si veda il bel sito sulle "macchine matematiche" all'indirizzo http://www.museo.unimo.it/theatrum/macchine/_00lab.htm.

riabile e cercando un sistema di due equazioni di cui tale equazione è la risultante. Ciò permette ad esempio di ottenere il seguente teorema.

Teorema 5.3. Fissiamo la parabola di equazioni $y = x^2$, allora si può risolvere
- qualsiasi equazione di secondo grado intersecando tale parabola con una retta
- qualsiasi equazione di terzo e quarto grado intersecando tale parabola con un'iperbole

Dim. Partiamo dall'equazione

$$ax^4+bx^3+cx^2+dx+e= 0 \qquad (5.1)$$

ed effettuiamo la sostituzione $y = x^2$. Otteniamo

$$ay^2+bxy+cy+dx+e = 0$$

Questo significa che (5.1) è la risultante del sistema

$$ay^2+bxy+cy+dx+e = 0$$
$$y = x^2.$$

La prima equazione è la parabola fissata, la seconda è un'iperbole. Ponendo $a = 0$ si ottiene un modo per risolvere le equazioni di terzo grado. Ponendo $a = b = 0$ otteniamo che una equazione di secondo grado si risolve intersecando la parabola $y = x^2$ con la retta $cy+dx+e = 0$. □

Esempio. Consideriamo l'equazione

$$x^4-x^3-3x^2-4 = 0. \qquad (5.2)$$

Ponendo $y = x^2$ otteniamo l'equazione $y^2-xy-3y-4 = 0$. Pertanto l'equazione è la risultante del sistema

$$y^2-xy-3y-4 = 0 \quad ; \quad y = x^2.$$

e quindi le radici si trovano intersecando le relative coniche.

Esistono molti altri modi di procedere nel trattare le equazioni. Ad esempio vale la seguente proposizione.

Teorema 5.4. Fissiamo l'iperbole di equazione $xy =1$, allora le equazioni di secondo e terzo grado si possono risolvere intersecando tale iperbole con una opportuna retta o, rispettivamente, con una opportuna parabola.

Dim. Consideriamo una equazione di terzo grado, non è restrittiva supporre che abbia la seguente forma

$ax^3+bx^2+cx -1 = 0.$ (5.3)

Possiamo mettere in evidenza x nei primi tre monomi

$x(ax^2+bx+c)-1 = 0$

e poi porre $y = ax^2+bx+c$. Sostituendo otteniamo $xy = 1$. Pertanto l'equazione (5.2) è la risultante del sistema

$xy = 1$; $y = ax^2+bx+c$.

Il caso di secondo grado si ottiene ponendo $a = 0$. □

Esempio. Consideriamo l'equazione

$x^3+2x^2-3x+1 = 0$

che riscriviamo nella forma $x(-x^2-2x+3) = 1$. Ponendo $y = -x^2-2x+3$ otteniamo $xy = 1$. Pertanto l'equazione è la risultante del sistema

$y = -x^2-2x+3$; $xy = 1$.

Problema. Costruire le soluzioni delle seguenti equazioni:

$x^4-2x^3+x-2 = 0$; $2x^3-x-14 = 0$.

Teorema 5.5. Le macchine che tracciano coniche permettono di risolvere i problemi della duplicazione del cubo e della trisezione dell'angolo.

Dim. Sia l la lunghezza dello spigolo di un cubo. Allora il volume di tale cubo è l^3. Dobbiamo costruire un segmento la cui lunghezza x sia tale che $x^3 = 2\cdot l^3$. Come abbiamo visto con Cartesio, non è difficile calcolare il prodotto di due segmenti e quindi costruire con riga è compasso un segmento di lunghezza $u = 2\cdot l^3$. La duplicazione del cubo si riconduce allora alla risoluzione di una semplice equazione di terzo grado $x^3 = u$. Per fare questo, posto $y = x^2$ e sostituendo si ottiene $x\cdot y = 2\cdot l^3 = u$. Quindi l'equazione $x^3 = u$ è la risultante del sistema

$y = x^2$

$x\cdot y = u$

Pertanto per duplicare il cubo di spigolo l dobbiamo tracciare una parabola, una iperbole e poi effettuare l'intersezione.

Anche la trisezione dell'angolo si riconduce alla soluzione di una equazione di terzo grado ma omettiamo la dimostrazione che richiederebbe un po' di trigonometria. □

6. Un potente calcolatore analogico: il regolo calcolatore

Nel 1614 Nepero introduce i logaritmi, grandezze tabulate che consentono di ricondurre le moltiplicazioni alle addizioni e le divisioni alle sottrazioni, semplificando grandemente i calcoli. Infatti ricordiamo che la funzione logaritmo (ad esempio in base 10) gode della seguente fondamentale proprietà

$Log(x \cdot y) = Log(x)+Log(y)$.

A sua volta tale equazione può essere riscritta al modo seguente

$x \cdot y = 10^{Log(x)+Log(y)}$.

Tale ovvia equazione può essere vista come un modo per ridurre la moltiplicazione all'addizione. Infatti per moltiplicare x con y

1. devo trasformare i numeri x ed y nei rispettivi logaritmi
2. devo eseguire l'addizione $Log(x)+Log(y)$
3. devo tornare indietro tramite la funzione inversa del logaritmo (cioè l'esponenziale).

Scrivendo delle tavole che permettono di associare ad ogni numero il relativo algoritmo e viceversa, il tutto si riduce ad una semplice addizione. Similmente, se si tiene conto dell'equazione

$Log(x/y) = Log(x)-Log(y)$.

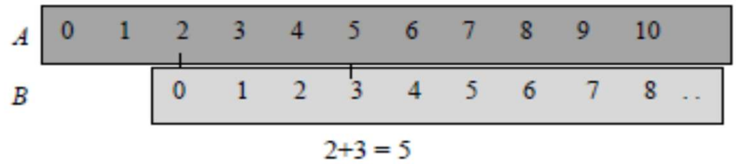
otteniamo la formula

$x/y = 10^{Log(x)-Log(y)}$.

Tale equazione permette di ridurre l'operazione di divisione ad una semplice sottrazione.

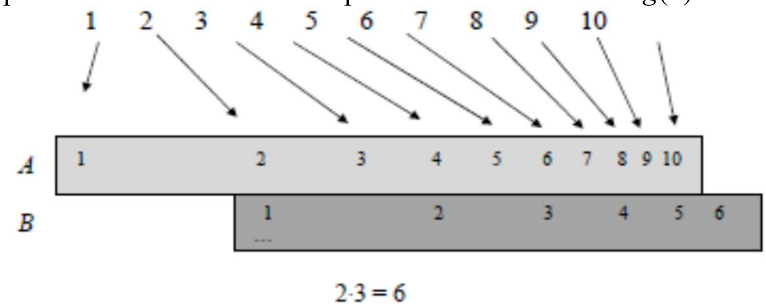
Le tavole dei logaritmi, soprattutto quelli in base 10 introdotte dall'inglese Henry Briggs (1561-1630), sono state utilizzate proficuamente fino ai giorni nostri.

Lo sviluppo naturale delle tavole logaritmiche è il *regolo calcolatore*. Questa macchina, è costituita da due listelli scorrevoli con tacche numerate. Per capire l'idea cominciamo con l'osservare che l'addizione di due numeri corrisponde al "fenomeno fisico" della traslazione di un righello rigido lungo una linea. Pertanto se vogliamo costruire una macchina per le addizioni (ad esempio per i numeri da 0 a 10) possiamo procedere al modo seguente. Costruiamo due regoli uguali con 11 tacche equidistanti numerate da 0 a 10:



Allora per sommare 2 a 3 è sufficiente portare la tacca 0 di *B* sulla tacca 2 di *A* e poi leggere (in *A*) in corrispondenza della tacca 3 (presente in *B*).

Poniamoci ora il problema di costruire un regolo capace di fare le moltiplicazioni, ad esempio di numeri compresi tra 1 e 10. A tale scopo traduciamo semplicemente “in modo fisico” il procedimento delle tavole dei logaritmi. Costruiamo infatti due listelli con una serie di “tacche” numerate da 1 a 10. Questa volta però la tacca *n*-esima viene posta ad una distanza *Log*(*n*) dalla

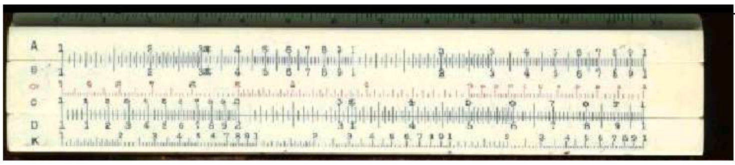


prima tacca. Ad esempio la tacca corrispondente al numero 2 viene posta ad una distanza *Log*(2) = 0.30 dalla prima tacca, quella corrispondente al numero 3 ad una distanza *Log*(3) = 0.47, Allora è evidente che per moltiplicare, ad esempio, 2 per 3:

- devo porre la prima tacca di *B* al posto 2 di *A*
- devo poi andare a leggere nel punto 3 di *B* il punto corrispondente in *A*.

Naturalmente ottengo il numero 6. Infatti il segmento 1-3 di *B* (di lunghezza *Log*(3)) ha effettuato una traslazione pari al segmento 1-2 (di lunghezza *Log*(2)) facendo raggiungere al punto tre una posizione pari a *Log*(2)+*Log*(3) = *Log*(6). Ma questa posizione corrisponde alla tacca del 6.

Se ci limitiamo ai numeri da 1 a 10 un tale tipo di calcolatore appare alquanto limitato perché può eseguire solo moltiplicazioni per numeri interi da 1 a 10. Se si vuole estendere tale tecnica di calcolo a numeri diversi basta prolungare i righelli oppure inserire altre tacche tra due successive.



I regoli, insieme alle tavole logaritmiche, sono state le calcolatrici più diffuse e fino a pochi anni fa non esisteva ingegnere o geometra che non ne possedesse uno. Sopra ho riprodotto un regolo calcolatore tascabile in cui il regolo scorrevole era inserito al centro del regolo fisso. Si producevano anche regoli da tavolo molto più grandi. Infatti è evidente che maggiore è la grandezza di un regolo maggiore è la sua precisione.

Una calcolatrice analogica. Sia la macchina riga+compasso che il regolo calcolatore sono un esempi di *calcolatori analogici*, macchine cioè in cui i calcoli numerici vengono svolti rappresentando "fisicamente" le grandezze in gioco (e quindi in modo approssimato) e traducendo i calcoli matematici in trasformazioni "fisiche" operate sulla macchina. In questo caso ci si riferisce al fenomeno della traslazione di regoli rigidi lungo una direzione fissata. Una caratteristica di tali tipi di macchine, che lavorano sul continuo, è quella della *interpolazione* e di un buon *controllo dell'errore*.

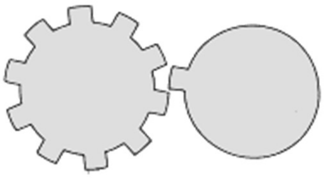
- **Interpolazione** significa che nel caso in cui i numeri da moltiplicare non sono segnati sul regolo, lo stesso può essere eseguito il calcolo portando le tacche in punti che approssimativamente corrispondono a tali numeri. In tale modo il risultato ottenuto è approssimativamente giusto.

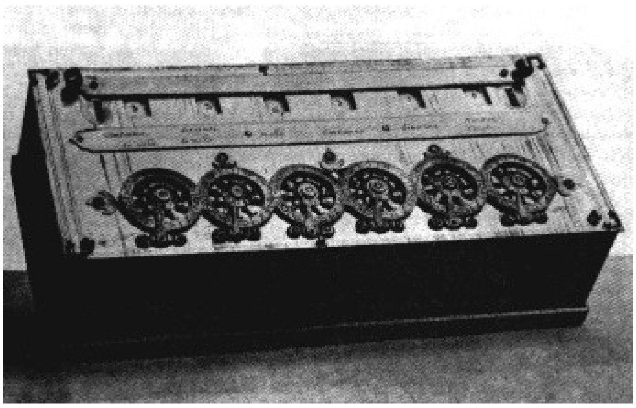
- **Controllo dell'errore** significa che se viene commesso un piccolo errore nel posizionare le tacche e nel far scorrere i regoli allora i risultati ottenuti risultano ancora adeguatamente esatti.

A questo genere di macchine si contrappongono le cosiddette *macchine digitali*, di cui sono un esempio notevole gli attuali computer elettronici. Tali macchine lavorano sul discreto, accettano solo dati precisi e non hanno un naturale meccanismo di interpolazione. Basta inoltre un minimo errore nell'introduzione dei dati (ad esempio la mancanza di una virgola oppure l'avere scritto una cifra in più) perché i risultati divengano completamente sbagliati. L'abaco è un esempio di macchina digitale.

7. Pascal e Leibniz: filosofi che fabbricano calcolatrici

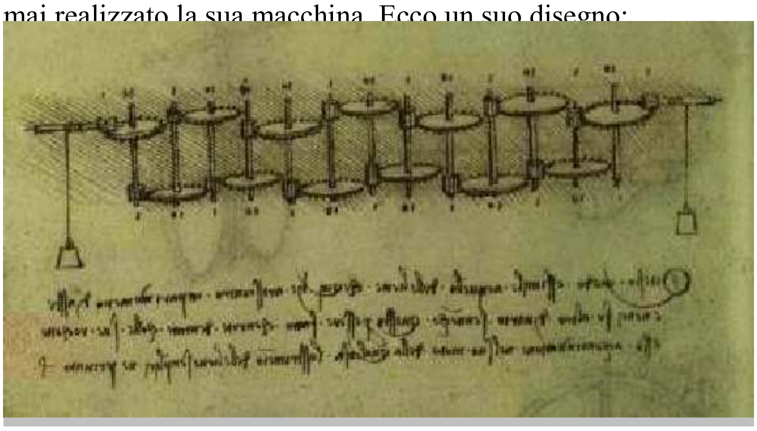
Nel corso del '600 vedono la luce anche le prime calcolatrici digitali meccaniche. Nel 1623 infatti Schickard, professore di matematica e astronomia a Tubigen realizzò un prototipo che tuttavia andò distrutto. La notizia di tale realizzazione ebbe pochissima diffusione. Ebbe invece successo una successiva realizzazione del matematico e filosofo francese Blaise Pascal il quale, volendo facilitare il padre nell'esecuzione di calcoli finanziari, nel 1642 realizza una calcolatrice meccanica, che in seguito verrà chiamata *pascalina.* Questa macchina, in grado di effettuare addizioni e sottrazioni, si basava sui movimenti di ruote dentate: l'addizione corrisponde ad effettuare rotazioni successive; e quando una ruota completa un giro fa avanzare di una posizione la ruota successiva realizzando l'operazione del riporto. La macchina era costituita da una serie di ruote dentate, una rappresentava le unità, la successiva le decine, poi le centinaia e così via (le ruote dentate erano 8). Sulla loro circonferenza erano scritti i numeri da 0 a 9. Era possibile eseguire addizioni e sottrazioni con riporto automatico fino ad otto cifre. Nel disegno, viene esemplificato il meccanismo che permette alla pascalina di riportare i numeri da una ruota a quella successiva. Ogni volta che la ruota di destra (le unità) compie un giro completo, il suo unico dente fa scattare di un'unità la ruota successiva posta a sinistra (le decine).



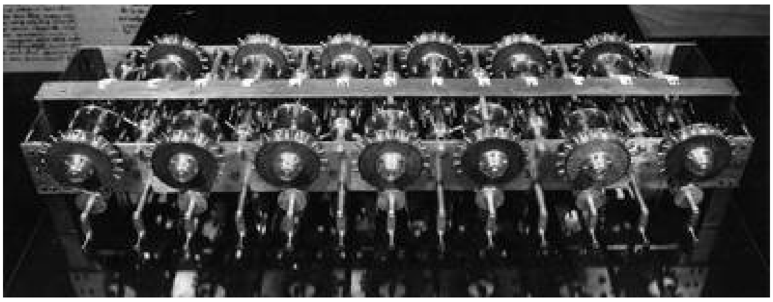


Sopra mostriamo una pascalina, sotto l'interno di una pascalina.

E' importante anche osservare che fin dal 1500 Leonardo da Vinci progettò una calcolatrice con una serie di ruote dentate opportunamente concatenate. Non sembra che Leonardo abbia mai realizzato la sua macchina. Ecco un suo disegno:



ed ecco una sua realizzazione moderna.



Ispirato dalla macchina di Pascal, un altro grande matematico e filosofo, il tedesco Gottfried Wilhelm Leibniz (1646-1716), realizza nel 1671 una calcolatrice meccanica in grado di effettuare le quattro operazioni e l'estrazione di radice. La moltiplicazione si riconduce a ripetute somme di uno dei fattori e di suoi multipli di 10, 100, ...

Uno degli ostacoli principali che si presentarono nella costruzione di calcolatrici meccaniche, consiste nella difficoltà di costruirne i componenti cosa che richiedeva una grande abilità e precisione e questo faceva salire molto il loro costo. In effetti nel '700 furono costruite varie macchine di questo genere, raffinate ma estremamente costose, le quali costituivano più degli oggetti ornamentali da regalare ad un re che degli strumenti di lavoro che potessero essere utilizzati da impiegati. Solo all'inizio dell'800 i progressi della meccanica verificatisi con lo sviluppo delle attività industriali hanno reso possibile produrre in serie meccanismi di elevata precisione e quindi di costruire calcolatrici commerciabili che furono utilizzate principalmente in ambito bancario. La prima di queste l'*Arithmometre* dovuto a Thomas de Colmar, un assicuratore-tecnologo, fu realizzata nel 1820 ed ebbe un buon successo commerciale fino alla fine del secolo.

Successo ancor maggiore si ebbe, a partire dalla fine dell' 800, con la realizzazione di calcolatrici elettromeccaniche che sono rimaste in uso fino a tempi recentissimi. La loro concezione non era diversa da quelle meccaniche ma semplicemente il movimento delle manovelle era alimentato da motorini ad energia elettrica cosa che aumentava notevolmente la velocità di calcolo.

8. Babbage, Ada Lovelace e la programmazione

Innovazioni fondamentali per le macchine di calcolo vennero proposte (e purtroppo realizzate solo in piccola parte) nella prima metà dell'800 dall'inglese Charles Babbage (1792-1871) che può considerarsi il grande precursore del calcolo automatico. Egli nel 1822 era riuscito a realizzare una macchina, chiamata *differential engine*. Se le macchine precedenti erano in grado al più di eseguire le operazioni aritmetiche di addizione, moltiplicazione, sottrazione, divisione ed estrazione di radice, tale macchina consentiva di calcolare valori di polinomi con 6 cifre decimali; a questa macchina si potevano richiedere operazioni diverse modificando l'assetto di alcune rotelle di controllo. Bisogna tenere presente che ogni funzione con opportune proprietà di continuità può essere sviluppata in serie di potenze e quindi approssimata tramite un opportuno polinomio. Pertanto la *differential engine* permetteva la costruzione di tavole numeriche in maniera rapida ed efficiente. Questo successo consentì a Babbage di ottenere dal governo inglese un ingente finanziamento per la costruzione di una macchina molto più complessa che doveva operare con 26 decimali. Questa realizzazione richiedeva la soluzione di moltissimi problemi meccanici che impegnarono pesantemente Babbage.

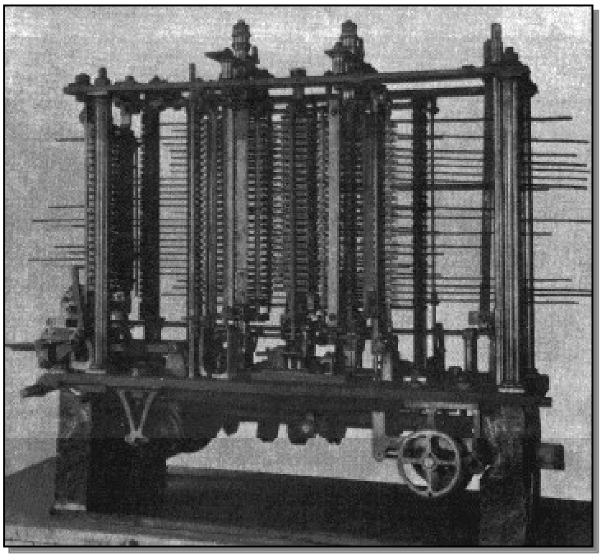


Figura 1: La macchina analitica di Babbage

Poi Babbage era portato ad apportare continui miglioramenti ai progetti e quindi a non concludere mai. Infatti nel 1830 Babbage fu indotto a progettare una macchina ancor più ambiziosa chiamata *analytical engine*, la quale, nelle intenzioni del suo ideatore, avrebbe dovuto avere molte caratteristiche del moderno computer[11]. Babbage ebbe l'idea di dotare il suo calcolatore di disposi-

[11]Intanto nel 1805 il francese Joseph Marie Jacquard (1752-1834) aveva realizzato il primo telaio automatico pilotato da schede perforate che consentivano di controllare il disegno delle stoffe e dei tappeti in produzione. Tale telaio può essere visto come una macchina programmabile e le schede possono essere viste come dei veri e propri "pro-

tivi che consentissero di servirsi di schede perforate per introdurre sia i dati da elaborare che istruzioni in grado di controllare la sequenza delle operazioni da effettuare ad ogni esecuzione. Per l'*analytical engine* infatti erano previsti registri nei quali memorizzare dati numerici separati dai dispositivi di calcolo; questi avrebbero dovuto essere governati da un controllo sequenziale dotato di possibilità di scelte e diramazioni.

La realizzazione dell'analytical engine, benché Babbage vi profondesse impegno e parte del suo patrimonio, incontrava grandi difficoltà. In effetti le limitazioni della tecnologia del tempo richiedevano di progettare e produrre quasi tutte le componenti ed il progetto era estremamente impegnativo, anche al livello del disegno dei pezzi.

La più importante collaboratrice di Babbage fu la figlia del poeta Lord Byron, Ada Augusta contessa di Lovelace (1815-1852), la quale si dedicò allo studio delle istruzioni da fornire all'*analytical engine* per calcoli impegnativi. Molti storici la considerano la prima programmatrice della storia. Alla fine il progetto della macchina analitica fu abbandonato in quanto il governo inglese non era disposto a proseguire nel finanziamento di un'impresa di cui non si vedeva la conclusione. Resta il fatto che in questa impresa sono state sviluppate alcune idee fondamentali per la progettazione di macchine per il calcolo automatico. Tali idee rimarranno inutilizzate e dimenticate fin poco prima della II guerra mondiale.

Babbage non completò mai la sua macchina. Gli ingranaggi e il vapore non erano all'altezza del compito che erano chiamati ad assolvere. Alla fine della macchina analitica fu costruito solo un piccolo modello che non funzionò mai.

9. Turing: logica matematica e teoria astratta dei calcolatori

Nella prima metà del novecento vi furono enormi progressi per i calcolatori, prima da un punto di vista teorico e poi da un punto di vista tecnologico. Un contributo teorico essenziale ad una definizione rigorosa della nozione "macchina che effettua calcoli" fu portato dalla logica matematica. Come vedremo nel prossimo

grammi" che determinavano il comportamento della macchina. Da notare che un bellissimo telaio di tale tipo, ancora perfettamente efficiente, si trova presso l'antica fabbrica di seta di San Leucio, vicino Caserta.

Alan Turing

volume, a partire dalla fine dell'800 si era sviluppato un importante filone di studi teso a fornire solidi fondamenti alla matematica e che era culminato nella proposta del grande matematico tedesco David Hilbert (1862-1943) di basare la matematica su sistemi logici formali. La matematica per Hilbert poteva essere identificata con la manipolazione di “stringhe” di simboli. Tale manipolazione doveva consentire di passare da alcune stringhe di simboli chiamate *assiomi di una teoria* ad altre stringhe di simboli chiamate *teoremi della teoria*. Una manipolazione che ovviamente doveva avvenire in accordo con certe regole e che corrisponde a quella che noi chiamiamo “dimostrazione matematica”. Coerentemente con tale punto di vista Hilbert pose la seguente questione:
- *esiste un algoritmo generale che consenta di decidere se una asserzione dell'aritmetica sia un teorema o meno?*

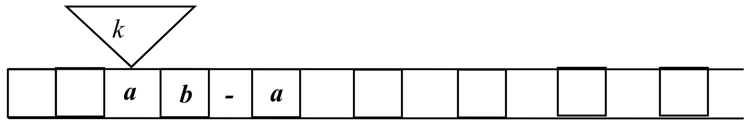
In risposta a tale problema il logico inglese Alan M. Turing nel 1935 scrive un fondamentale articolo dal titolo *“On computable numbers, with an application to the Entschaidungsproblem”* in cui dimostra la non esistenza di un tale algoritmo.

Per potere dimostrare un tale teorema ovviamente era prima necessario definire in modo rigoroso che cosa si debba intendere per “algoritmo”. Infatti mentre l'esistenza di un algoritmo può essere provata tramite una semplice esibizione di un processo che intuitivamente viene accettato da tutti come algoritmo, la situazione per provarne la non esistenza impone una precisa definizione della classe degli algoritmi.

A tale scopo Turing procede al modo seguente:

- definisce il concetto di algoritmo introducendo una nozione precisa e plausibile di "macchina che esegue un calcolo"
- mostra che nessuna delle macchine che entrano nella sua definizione è in grado di risolvere il problema suddetto.

Le “macchine” di Turing costituiscono una nozione completamente teorica e formale. Egli immagina marchingegni con testina di scrittura/lettura su di un nastro "potenzialmente" illimitato da entrambi i lati. Sul nastro possono essere scritti vari simboli di un dato alfabeto e la macchina può assumere vari stati. Ad ogni istante la macchina si trova in uno stato e la testina è posata su un punto del nastro dove legge quello che vi è scritto.



Il disegno rappresenta una macchina di Turing che è nello stato *k* e che legge il simbolo *a* scritto sul nastro. La macchina può fare poche cose e precisamente:

spostare la testina a destra, a sinistra, lasciarla ferma, stampare un carattere al posto di quello esistente.

Esponiamo più in dettaglio la definizione proposta da Turing.

Definizione 9.1. Una *macchina di Turing* è una struttura matematica del tipo $M = <A, -, K, k_0>$ dove:

A è un insieme finito chiamato *alfabeto*
\- è un carattere speciale, interpretato come *spazio bianco*
K è un insieme finito detto *insieme degli stati*
k_0 è uno stato detto *stato iniziale*

L'input viene scritto sul nastro sotto forma di una parola nell'alfabeto A cioè una sequenza di lettere dell'alfabeto. Il comportamento della macchina, che inizialmente è posta nello stato iniziale dipende in modo deterministico dal tipo di programma utilizzato.

Definizione 9.2. Un *programma* per una macchina di Turing è una funzione parziale $\delta : K\times(A\cup\{-\})\rightarrow K\times(A\cup\{-\})\times\{d,s,i\}$ che chiamiamo *funzione di transizione.*[12]

Interpretiamo i simboli d, s, i come "spostarsi a destra", "spostarsi a sinistra" e "stare fermi". Precisamente, per ogni $k, k' \in K$ e $c, c' \in A\cup\{-\}$ $\delta(k,c) = (k',c',d)$ significa che se la testina legge c e la macchina è nello stato k allora la testina:

1. passa dallo stato k allo stato k',

[12] Da notare che spesso viene chiamata macchina di Turing una struttura del tipo $<A, -, K, k_0, \delta>$ cioè quella che noi abbiamo chiamato macchina di Turing più un programma

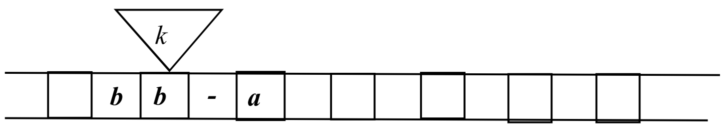
2. sostituisce c con c',
3. si sposta di un passo verso destra,

- $\delta(k,c) = (k',c',s)$ significa che se la testina legge c e la macchina è nello stato k allora la testina:
   1. passa dallo stato k allo stato k',
   2. sostituisce c con c',
   3. si sposta di un passo verso sinistra,
- $\delta(k,c) = (k',c',i)$ significa che se la testina legge c e la macchina è nello stato k allora la testina:
   1. passa dallo stato k allo stato k',
   2. sostituisce c con c',
   3. rimane ferma.

Per eseguire un calcolo scriviamo l'input come una parola x scritta sul nastro interpretando il simbolo – come casella vuota. Inoltre posizioniamo la testina in modo che guardi il primo simbolo diverso da -.

Definizione 9.3. Chiamiamo *configurazione istantanea* una parola del tipo $c_1...c_nkc_{n+1}...c_m$ dove $c_1,...,c_m$ sono caratteri e k è uno stato. Chiamiamo *configurazione iniziale dato l'input x* la configurazione k_0x.

La configurazione $c_1...c_nkc_{n+1}...c_m$ rappresenta la situazione in cui sul nastro è scritta la parola $c_1...c_m$ la testina è nello stato k ed è posizionata sul simbolo c_{n+1}. La figura che abbiamo indicato sopra corrisponde all'input *ab-a* (input che viene identificato con la coppia (ab,a)) e quindi alla configurazione $k_0ab\text{-}a$. Successivamente la macchina comincia ad agire in accordo con quanto è imposto dalla funzione di transizione. Ad esempio, se $\delta(k_0,a) = (k,b,d)$ allora la macchina dopo avere sostituito a con b passa nella casella a destra ed assume lo stato k (si veda la figura successiva)



Si perviene allora ad una configurazione che possiamo denotare con *bkb-a*.

Definizione 9.4. Consideriamo una macchina di Turing con un dato programma. Allora, dato un input $x \in A^+$, facciamo partire la macchina con la configurazione iniziale k_0x e lasciamo che la macchina passi alle configurazioni successive. Se nello stato k legge un simbolo $\underline{x}$ ma la funzione di transizione non è definita nella coppia $(k,\underline{x})$ facciamo fermare la macchina e la sequenza di simboli in A che sono rimasti sul nastro è l'output.

Se la macchina si ferma si dice anche che *converge* altrimenti si dice che *diverge.* In base a tale definizione una macchina di Turing calcola una funzione parziale di A^+ in A^+.

Definizione 9.5. Dato un alfabeto finito A, una funzione parziale di A^+ in A^+ è detta *Turing calcolabile*, se esiste una programma che la calcola in accordo con quanto detto nella definizione 9.4.

Se vogliamo definire la calcolabilità per funzioni di più variabili è sufficiente rappresentare le ennuple di elementi di $A^+ \times A^+$ con parole nel linguaggio $(A \cup \{-\})$. Ad esempio rappresentare una terna (x,y,z) di elementi di A^+ con la parola x-y-z.

10. Macchine a registri e tesi di Church

I moderni computer digitali sono molto lontani dall'assomigliare ad una macchina di Turing. Un modello matematico più vicino a tali computer è fornito dalla nozione di macchina a registri. Supponiamo di volere scrivere un programma per calcolare la somma di due numeri. Tale programma potrebbe essere costituito dalle seguenti cinque istruzioni:

1. poni x ed y uguali ai numeri da addizionare e vai a 2
2. poni $s := x$, $c := 0$ e vai a 3
3. poni $s := s+1$; $c := c+1$ e vai a 4
4. SE $c = y$ ALLORA vai a 5 ALTRIMENTI vai a 3
5. scrivi s.

Similmente un programma per la moltiplicazione potrebbe essere:

1. poni x ed y uguali ai numeri da moltiplicare e vai a 2
2. poni $p := x$, $c := 1$ e vai a 3
3. poni $p := p+x$; $c := c+1$ e vai a 4
4. SE $c = y$ ALLORA vai a 5 ALTRIMENTI vai a 3
5. scrivi p

In entrambi i casi le istruzioni 3 e 4 vengono ripetute fino a quando il *"contatore" c* non sia stato incrementato di una unità un numero di volte pari ad y. Appena c raggiunge y viene fornito l'output y. Chiamiamo *programma* una successione finita di istruzioni di questo tipo che supponiamo iniziare con una istruzione input e terminare con una istruzione di output. Ciascuna istruzione inizia con numero di indirizzo e finisce con l'indirizzo dell'istruzione da eseguire successivamente (con esclusione dell'istruzione di output). Un "hardware" per potere eseguire istruzioni di tale tipo deve consentire di effettuare le seguenti operazioni.

1. Devono essere memorizzati i valori delle variabili x, y, s, c e p. Questo può essere fatto tramite opportuni registri di memoria dove scrivere dei numeri naturali.
2. Deve essere possibile la modifica del valore delle variabili tramite alcune operazioni elementari. Nel nostro caso porre nel registro di memoria c il numero 0, aggiungere uno al registro di memoria s, …
3. Deve essere possibile un controllo dello stato assunto per poter eseguire istruzioni condizionate del tipo SE ... ALLORA ... ALTRIMENTI . . . (nel nostro caso controllare se $c = y$).

Ad ogni variabile corrisponde un registro di memoria e se si utilizzano m variabili abbiamo bisogno di m registri di memoria. Lo stato della macchine è dato dal contenuto dei registri e quindi da una m-pla di numeri naturali. Il calcolatore può quindi assumere come stato un qualsiasi elemento dell'insieme N^m. Possiamo immaginare che inizialmente i registri siano azzerati e che, dopo l'inserimento dell'output, la macchina esegui il programma passando attraverso una serie di stati. La macchina si ferma quando viene raggiunta l'istruzione di output. E' evidente che cosa si debba intendere per funzione di una o più variabili calcolata da un programma per macchine a registri.

Dobbiamo ora confrontare la nozione di calcolabilità fornita dalle macchine di Turing con quella fornita dalle macchine a registri. Una differenza consiste nel fatto che una macchina di Turing può calcolare funzioni in un insieme A^+ di parole mentre una macchina a registri lavora sui naturali.

Questa differenza è solo apparente. Infatti poiché i numeri naturali possono essere rappresentati da parole in una data base le macchine di Turing forniscono una nozione di calcolabilità anche su tale tipo di elementi. D'altra parte il fatto che esse possano lavorare sull'insieme delle parole di un qualunque alfabeto non costituisce una reale differenza. Questo perché quando scriviamo una parola, ad esempio sulla tastiera di un computer, questa parola si trasforma immediatamente in un numero naturale. Ciò si esprime dicendo che l'insieme delle parole su di un linguaggio finito è *codificabile*. Ad esempio, supponendo di rappresentare i numeri naturali in base 2, possiamo associare ad ogni lettera dell'alfabeto una sequenza di 0 ed 1, ad esempio di lunghezza otto. Ad esempio posso decidere di associare:

- alla lettera *a* il numero 10000000
- alla lettera *b* il numero 10000001
- alla lettera *c* il numero 10000010
- alla lettera *d* il numero 10000011
- alla lettera *e* il numero 10000100

. . .

Successivamente ad ogni parola associamo la sequenza ottenuta sostituendo a ciascuna lettera il codice corrispondente. Ad esempio alla parola *bacca* associo il numero, scritto in base due,

1000000110000000**10000000**10000001**10000000**

alla parola *ada* associo il numero

10000000**10000011****10000000**

alla parola *cade* associo il numero

1000001010000000**10000011**10000100

e così via.

Una volta messo in evidenza che non esiste differenza sul tipo di elementi (numeri o parole) trattati dalle macchine a registri e le macchine di Turing, vale il seguente teorema di cui omettiamo la dimostrazione che mostra come le due corrispondenti nozioni di computabilità coincidono.

Teorema 10.1. Le funzioni calcolabili da una macchina a registri sono tutte e sole le funzioni calcolabili da una macchina di Turing.

Tale teorema suggerisce il fatto che la nozione proposta da Turing di funzione computabile sia completamente adeguata. Questa ipotesi è confermata da numerosi altro teoremi simili. Infatti:

1. Dopo Turing sono state proposte numerose definizioni di

"macchina che effettua calcoli" o di "algoritmo". Per tutte è stato dimostrata l'equivalenza con le macchine di Turing.

2. Tutte le volte che è stato proposta un linguaggio di programmazione si è visto che computa esattamente le stesse funzioni calcolabili dalle macchine di Turing

3. Fino ad ora non è mai stata trovata una funzione (intuitivamente) calcolabile che non sia calcolabile tramite una opportuna macchina di Turing.

Questi fatti hanno portato al seguente convincimento che va sotto il nome di *"Tesi di Church"* :

> *la definizione di Turing di funzione calcolabile è perfettamente adeguata a rappresentare la nozione di funzione computabile.*

Tale tesi ci permette di parlare di *funzioni calcolabili* senza specificare lo strumento con cui vengono calcolate.

11. Teoremi limitativi per i calcolatori

Non è difficile vedere che la maggioranza delle funzioni di N in N non è computabile. La dimostrazione si basa su qualche nozione elementare di teoria degli insiemi. Per prima cosa osserviamo che ad ogni programma è possibile assegnare un numero di codice. Basta vedere un programma come un'unica parola ed applicare il procedimento di codifica indicato nel paragrafo precedente. Indicheremo allora con

$\pi_1, \pi_2, \ldots, \pi_i, \ldots$

i programma di codice 1, 2, …, i, …, rispettivamente e con

$f_1, f_2, \ldots, f_i$ ...

le rispettive funzioni calcolate.

Proposizione 11.1. Esistono funzioni che nessun calcolatore potrà mai calcolare. Precisamente l'insieme delle funzioni computabili ha la potenza del numerabile mentre l'insieme delle funzioni non computabili ha la potenza del continuo.

Dim. Abbiamo visto che ad ogni funzione computabile è possibile associare un numero di codice. Pertanto l'insieme delle funzioni computabili ha la potenza del numerabile. D'altra parte l'insieme delle funzioni parziali di N in N ha la potenza del continuo. Quindi l'insieme delle funzioni non computabili deve avere la potenza del continuo. □

La nozione di funzione calcolabile permette di dare la seguente fondamentale nozione.

Definizione 11.2. Chiamiamo *decidibile* un sottoinsieme S di N (o di un insieme codificabile) la cui funzione caratteristica sia calcolabile.

Allora un insieme S è decidibile se esiste un programma di una macchina a registri capace di dire se un elemento x appartiene ad S (output 1) o non ci appartiene (output 0). In proposito vale la seguente proposizione che si prova come la proposizione 11.1.

Proposizione 11.3. Esistono insiemi non decidibili, cioè insiemi per i quali il problema dell'appartenenza non potra essere mai risolto da un calcolatore. Precisamente mentre la classe degli insiemi decidibili ha la potenza del numerabile, la classe degli insiemi indecidibili ha la potenza del continuo.

Tale proposizione prova che esiste un insieme indecidibile ma non ci fornisce un esempio diretto. Tale esempio è dato dal seguente teorema, di cui non forniamo la dimostrazione, che fornisce una risposta negativa al problema posto da Hilbert.

Teorema 11.4. (Indecidibilità dell'aritmetica). L'insieme (dei numeri di codice) delle espressioni vere dell'aritmetica non è decidibile. In altre parole, non esiste un programma tale che, dato come input una asserzione dell'aritmetica, fornisca output 1 se l'asserzione è vera, 0 se è falsa.

Si potrebbe osservare che il teorema non si riferisce ad un insieme di numeri ma ad un insieme di asserzioni. Tuttavia se si pensa che ad ogni asserzione in un linguaggio può essere associato un opportuno numero di codice, considerare un insieme di asserzioni non è diverso dal considerare un insieme di numeri naturali.

E' interessante osservare che altri teoremi limitativi come questo li abbiamo trovati per la macchina "riga e compasso". Tuttavia abbiamo visto che si possono superare i limiti di tale macchina introducendo macchine che tracciano altre curve. Nel caso delle macchine di Turing la questione è più complicata. Infatti se si accetta la tesi di Church il teorema ora enunciato si estende a qualunque tipo di macchina presente o futura.

Teorema 11.5. Se si accetta la tesi di Church non potrà mai essere costruita una macchina o scritto un programma per cui, dato come input una asserzione dell'aritmetica, si abbia come output 1 se l'asserzione è vera, 0 se è falsa.

Di cose che nessuna macchina potrà mai fare (problemi indecidibili) ne esistono anche molte altre. Forniamo un esempio, che va sotto il nome di "problema della fermata".

Abbiamo visto che quando inseriamo un input sul nastro di una macchina e facciamo partire la macchina non è detto che la macchina si fermi. La stessa cosa avviene per le macchine a registri dove il "rimbalzo" da una istruzione all'altra potrebbe non-fermarsi mai.

La cosa potrebbe sembrare di scarso interesse ma coincide con un fenomeno di cui tutti gli utilizzatori di computer hanno una triste esperienza. Se scriviamo qualcosa con la tastiera (input) a volte vediamo che il computer pur lavorando alacremente non si decide a fermarsi ed a dare l'output desiderato. Si pone pertanto il problema se sia il caso di spegnere il computer (perdendo magari il lavoro già fatto) oppure no. Evidentemente sarebbe utilissimo avere un programma che eviti una tale difficoltà, e che quando inseriamo un input che crea problemi di non fermata ci avverti dicendo qualcosa come "attenzione: input pericoloso !". Purtroppo la cosa non è possibile come dimostra il seguente teorema.

Teorema 11.6. (Problema della fermata). L'insieme

$$K_0 = \{(i,j) \in N \times N : \pi_i \text{ converge in } j\} \qquad (11.1)$$

non è decidibile. In altre parole non esiste un programma capace di controllare se, dato un qualsiasi programma π_i ed un input j, π_i si fermi o meno in corrispondenza di j.

Dim. Cominciamo col dimostrare che l'insieme

$$K = \{i \in N : \pi_i \text{ converge in } i\} \qquad (11.2)$$

non è decidibile. Infatti supponiamo per assurdo che la sua funzione caratteristica c_K sia computabile. Allora esiste un programma π per calcolare c_K. Aggiungiamo alle istruzioni in π ulteriori istruzioni in modo da avere il seguente programma:

$$\left.\begin{array}{l} 1.\ \dots \\ \quad\dots \\ n.\ \dots \end{array}\right\} \quad \text{<programma } \pi \text{ per calcolare } c_K\text{>}$$

n+1. input (i) vai a $n+2$
n+2. se $c_K(i) = 1$ allora vai a n+2 altrimenti vai a n+3
n+3. Stampa 1

E' evidente che tale programma converge se e solo se $c_K(i) = 0$ pertanto, se d è il suo numero di codice

$\pi_d = \pi$ converge in $d \Leftrightarrow c_K(d) = 0 \;\Leftrightarrow\; \pi_d$ non converge in d.

L'assurdo a cui siamo pervenuti ci dice che l'ipotesi di esistenza del programma π è falsa.

Dalla indecidibilità di K segue l'indecidibilità di K_0 in quanto in caso contrario per potere decidere se $i\in K$ basterebbe vedere se $(i,i)\in K_0$. □

La nozione di decidibilità può essere indebolita al modo seguente.

Definizione 11.7. Un sottoinsieme S di N è detto *effettivamente enumerabile* se è l'insieme vuoto oppure se esiste una funzione totale computabile suriettiva $h : N\rightarrow S$.

Ad esempio l'insieme dei quadrati perfetti $\{n^2 : n\in N\}$ essendo, per definizione, il codominio della funzione n^2, è effettivamente numerabile. Allora un insieme non vuoto è effettivamente enumerabile se i suoi elementi si possono "produrre" uno dopo l'altro in una successione $h(1)$, $h(2)$, ….

Proposizione 11.8. Ogni insieme decidibile S è anche effettivamente enumerabile.

Dim. Infatti S o è vuoto e quindi effettivamente enumerabile per definizione oppure contiene un elemento $\underline{s}$. In questo caso basta definire la funzione h ponendo $h(n) = n$ se $n\in S$ e $h(n) = \underline{s}$ in caso contrario. Poiché S è decidibile tale funzione è effettivamente calcolabile. Poiché è evidente che il codominio di h è S, S risulta essere effettivamente enumerabile. □

Teorema 11.9. Il sottoinsieme K_0 relativo al problema della fermata è effettivamente enumerabile (pur non essendo decidibile).

Dim. Dato un programma π possiamo supporre di modificarlo introducendo una nuova variabile c, chiamata *contatore*, ponendo $c := 0$ inizialmente ed aggiungendo ad ogni istruzione il comando $c := c+1$. Se eseguo un calcolo con input i allora alla fine c avrà assunto un valore che misura il numero di volte in cui una istruzione è stata eseguita. In questo caso diciamo che π converge in c passi. Detto questo, definiamo h al modo seguente. Sia $(\underline{i}, \underline{j})$) un elemento fissato di K_0, poniamo allora

$h(n) = (i,j)$ se scomposto n in fattori primi $n = 2^i \cdot 3^j \cdot 5^c \cdot \ldots$ il programma π_i con input j converge in c passi

$h(n) = (a,b)$ altrimenti.

E' evidente che h è computabile e che il suo codominio è K_0. □

Finiamo questo paragrafo con un un teorema legato al processo di integrazione (non numerica). Consideriamo un linguaggio con le usuali operazioni algebriche e con nomi per le funzioni elementari (seno, coseno, esponenziale, …). Indichiamo con *El* l'insieme dei temini di tale linguaggio. Ad esempio in *El* troveremo parole come $seno(x)+3coseno(x^2)$, $\sqrt{seno(x)+5}$….

Quando in un corso di analisi si studia la derivazione, in realtà si studia come possa essere calcolata la funzione $D : El \rightarrow El$ che a partire da un termine denotante una funzione permette di ottenere un altro termine denotante la relativa derivata. L'integrazione indefinita può, in un certo senso, essere vista come la funzione inversa (a meno di una costante). Un termine t' si dice *integrabile elementarmente* se appartiene al codominio di D cioè se esiste un termine t tale che $D(t) = t'$. Detto più semplicemente, una funzione è integrabile elementarmente se la sua primitiva può essere definita a partire dalle funzioni elementari. Possiamo ora enunciare il seguente teorema che a me sembra di un interesse straordinario.

Teorema 11.10. L'insieme delle formule integrabili elementarmente è effettivamente enumerabile ma non decidibile.

Dim. La funzione $D : El \rightarrow El$ è effettivamente computabile, infatti le usuali regole di derivazione costituiscono un algoritmo che permette il calcolo della derivata per ricorsione. La base della ricorsione è l'elenco delle derivate delle funzioni elementari. Ciò prova che l'insieme $D(El)$ delle funzioni integrabili elementarmente è effettivamente enumerabile. La parte difficile del teorema, cioè la non decidibilità, viene omessa. □

Il fatto che D sia effettivamente computabile non è sorprendente in quanto il calcolo di D si effettua tramite un semplice processo di ricorsione. Il fatto che $D(El)$ non sia decidibile significa che non esiste un metodo generale per verificare se una funzione, descritta tramite funzioni elementari, ammette una primitiva che sia ancora descrivibile tramite le funzioni elementari.

12. E' veramente possibile addizionare due numeri reali?
Utilizzando la nozione di funzione computabile è possibile analizzare e rivedere tutte le nozioni della matematica da un punto di vista "costruttivista". Ad esempio possiamo chiamare "computabile" un numero reale la cui espansione decimale è una funzione computabile. Tutti i numeri normalmente utilizzati dai matematici sono computabili. Ad esempio l'algoritmo per calcolare l'espansione decimale della radice di due ci dice che questo numero è computabile. Tuttavia ancora una volta si vede che la maggioranza dei numeri reali risulta non essere computabile.

Un problema interessante è se le operazioni di addizione e di moltiplicazione siano effettivamente eseguibili anche se ci si limita all'ambito dei numeri computabili. Usualmente viene dato per scontato il seguente principio:

date le espansioni decimali di due numeri x ed y è sempre possibile ottenere l'espansione decimale di $x+y$.

Siamo sicuri che esso valga ? Dal punto di vista della matematica classica non esistono problemi. Poiché ogni numero reale può essere scritto in forma decimale, non ci sono dubbi che anche $x+y$ può essere scritto in forma decimale. Tuttavia, se si assume un punto di vista "più concreto" o, se si vuole "costruttivista", la risposta, contrariamente a quanto la matematica classica ci ha abituato a pensare, è *negativa.* La cosa è stata dimostrata da B. H. Mayoh nell'ambito della teoria della ricorsività.

Teorema 12.1. Nell'insieme delle espansioni decimali dei numeri reali non è effettivamente eseguibile la moltiplicazione. Ne segue che non è effettivamente eseguibile nemmeno l'operazione di addizione.

Dim. Mi limito a dimostrare l'impossibilità di effettuare la moltiplicazione per tre in quanto ciò implica anche l'impossibilità di effettuare l'addizione. Infatti, poiché $3x = (x+x)+x$, se esistesse un procedimento effettivo per l'addizione, allora esisterebbe anche un procedimento effettivo per la moltiplicazione per tre.
In proposito invece di procedere in una dimostrazione rigorosa in cui si applicano le nozioni dei paragrafi precedenti, provo a proporre prima due dimostrazioni intuitive che potrebbero essere comprese anche da studenti digiuni di teoria della computabilità.

Prima "dimostrazione"
Immaginiamo di rivolgerci ad uno studente, Carlo, al modo seguente:

> *scommettiamo che se ti fornisco l' espansione decimale di un numero x non sei nemmeno in grado di dirmi quale è la parte intera dell'espansione decimale di* $3x$ *?*

Una volta accettata la scommessa si procede a chiarire meglio le regole del gioco. Io mi impegno a comunicare la parte intera di x e poi una dopo l'altra tutte le cifre decimali. Carlo deve, prima o poi, cominciare a fornire una risposta dicendo quale è la parte intera e poi man mano le cifre decimali di $3x$. Concordiamo di non considerare mai il periodo 9, visto che tale periodo si può sempre evitare. Supponiamo allora che io cominci a dire a Carlo "*il numero x ha la parte intera uguale a 0, la prima cifra è 3, la seconda 3, ...*"
Prima o poi Carlo è tenuto a dire quale è la parte intera di $3x$. Tuttavia:
- se dopo avere comunicato la cifra m-esima Carlo dice che la parte intera di $3x$ è 0 rispondo che ha sbagliato perché tutte le successive cifre di x sono uguali a 3 e quindi, essendo $x = 1/3$, il numero $3x$ è uguale ad 1
- se Carlo afferma che tale parte intera è 1 allora dico ancora che ha sbagliato perché tutte le cifre successive alla cifra m-esima sono uguali a 0 e quindi, essendo $x<1/3$, abbiamo che $3x <1$.

Naturalmente non tutti sarebbero d'accordo a ritenere che questa sia una dimostrazione ed in effetti ha un po' il sapore di un imbroglio in quanto dovrei fissare prima quale è il numero magari scrivendolo su di un foglietto da consegnare ad un notaio. Ma la cosa non presenta problemi. Sul foglietto posso dichiarare quale è il numero reale e posso perfino fare leggere a Carlo il contenuto del foglietto prima ancora di iniziare l'esperimento. Basta che io scriva "x è il numero $0{,}x_1x_2\ldots x_i\ldots$ la cui parte intera è 0 ed inoltre

$x_i = 3$ se Carlo all'istante i-esimo non ha ancora fornito la prima cifra decimale o ha risposto 0

$x_i = 0$ se Carlo all'istante i-esimo ha già fornito come risposta 1.

Nonostante questa esplicita definizione di x, Carlo non ha speranze di dire in modo corretto quale è la parte intera di $3x$.

Seconda "dimostrazione"

Certo la definizione ora data di x è …"poco matematica" e la dimostrazione ha il sapore di uno dei soliti trucchi usati dai logici per dimostrare cose strane. Possiamo allora trovare una definizione migliore di x ed una diversa dimostrazione utilizzando tecniche della matematica intuizionista. Ad esempio consideriamo il problema se nella espansione di π compaia la sequenza delle cifre 0123456789 oppure no. Allo stato attuale delle conoscenze non sappiamo fornire una risposta in proposito. Indichiamo con $A(i)$ l'affermazione: « i è il primo intero tale che tra le prime i cifre di π compare 012...9 » e consideriamo il numero reale $x = 0{,}x_1 \ldots x_i \ldots$ tale che:

$x_i = 0$ se $A(i)$ è vera

$x_i = 3$ altrimenti.

Poiché conosciamo un algoritmo per l'espansione decimale di π, abbiamo definito x tramite un esplicito algoritmo matematico per generare x. Sono possibili ora due casi:

1° *caso.* - $A(i)$ è sempre falsa poiché la sequenza 0123456789 non compare mai e quindi $x = 1/3$ e $3x = 1$

2° *caso.* - $A(i)$ è vera per un numero i e quindi $x<1/3$ e $3x<1$

Pertanto la parte intera di $3x$ sarà uguale ad 1 nel primo caso ed a 0 altrimenti. Poiché non siamo in grado di decidere se vale il primo o il secondo caso, non siamo in grado di determinare la parte intera di $3x$.

Terza dimostrazione
Chi non ama l'intuizionismo potrebbe osservare che tale dimostrazione è troppo legata al fatto storico che fino ad ora il problema della presenza della sequenza 0123456789 non è stato ancora risolto. Basta allora che qualche matematico geniale risolva la questione perché la dimostrazione non sia più valida. Propongo allora quest'ultima dimostrazione che si basa sulla teoria della computabilità esposta nel paragrafo precedente. Indichiamo con S un sottoinsieme di N che non sia decidibile ma che sia effettivamente enumerabile, cioè che sia il codominio di una funzione computabile totale $h : N \to N$. Per ogni $n \in N$ consideriamo il numero reale $x = 0{,}x_1...x_i...$ definito ponendo

$x_i = 3$ se $h(i) \neq n$

$x_i = 0$ altrimenti.

Allora l'espansione decimale di x avrà tutte le cifre uguali a 3 nel caso in cui $n \notin S$ e tutte le cifre uguali a 3 tranne una uguale a 0 nel caso in cui $n \in S$. Ciò comporta che

- se $n \in S$, allora essendo $x<1/3$, deve essere $3x<1$
- se $n \notin S$ allora essendo $x = 1/3$ deve essere $3x = 1$.

Equivalentemente,

$n \in S \Leftrightarrow$ la parte intera di $3x$ è 0.

Pertanto, se fossi in grado di fornire la parte intera di $3x$ sarei in grado di decidere anche se n appartiene ad S oppure no. Questo contraddice la supposta non decidibilità di S. □

In apparente contrasto con quanto ora dimostrato vi è la seguente ovvia proposizione.

Proposizione 12.2. Nell'insieme delle successioni di Cauchy l' operazione di somma può essere effettivamente eseguita.

Dim. E' evidente che se ho un algoritmo per stampare i valori di due successioni di Cauchy $(r_n)_{n\in N}$ e $(q_n)_{n\in N}$ di razionali, allora ho anche un algoritmo per stampare i valori della loro somma $(r_n)_{n\in N}$ $+(q_n)_{n\in N} = (r_n+q_n)_{n\in N}$. □

Si potrebbe obiettare che ogni espansione decimale è anche una successione di Cauchy e che quindi tale proposizione, in contrasto con il teorema di Mayoh, comporterebbe la possibilità di effettuare la somma di due espansioni decimali. Detto in modo più preciso, dati due numeri reali computabili $x = x_0,x_1...x_i...$ e $y = y_0,y_1 ...y_i ...$ possiamo considerare le successioni $(r_n)_{n\in N}$ e $(q_n)_{n\in N}$ ottenute ponendo $r_n = x_0,x_1...x_n$ e $q_n = y_0,y_1y_2...y_n$. Otteniamo due successioni di Cauchy di razionali effettivamente computabili e quindi $(r_n+q_n)_{n\in N}$ è una successione di Cauchy effettivamente computabile che rappresenta $x+y$. Purtroppo, come abbiamo visto nella dimostrazione della proposizione 10.1, non è detto che si sia capaci di trasformare tale successione in una espansione decimale effettivamente computabile. Infatti, pur essendo r_n+q_n un numero razionale scritto in forma decimale finita, non è detto che costituisca la troncata di livello n della espansione di $x+y$.

Naturalmente non si esclude che possano esistere particolari classi di numeri reali per cui l'addizione sia effettivamente eseguibile. Ad esempio ciò accade per l'insieme dei reali che sono razionali i cui algoritmi di espansione decimale siano dati tramite il periodo e l'antiperiodo. Quello che si esclude è che possa esistere un unico procedimento valido per ogni possibile coppia di algoritmi per espansioni decimali di numeri reali.

Oltre che per l'addizione, valgono analoghe considerazioni anche per le altre operazioni come mostra la seguente proposizione.

Proposizione 12.3. Nell'insieme delle espansioni decimali dei numeri reali non sono eseguibili in modo effettivo le operazioni di sottrazione, divisione, moltiplicazione ed estrazione di radice.

Succedono anche altri fatti abbastanza strani. Ad esempio mentre abbiamo visto che la moltiplicazione per 3 non è eseguibile, quella per 2 lo è. Se si considerano rappresentazioni dei numeri in base 3 vale il viceversa. Inoltre il passaggio da una base p ad una base q è possibile (in modo effettivo) solo quando q divide una potenza di p.

Quanto detto in questo paragrafo mostra che il definire i numeri reali identificandoli con le espansioni decimali è una cosa comoda ma alquanto criticabile e che la migliore definizione di numero reale è probabilmente quella proposta da Cantor tramite le successioni di Cauchy. D'altra parte tale definizione corrisponde al modo usuale di utilizzare i numeri reali.

CAPITOLO 5
DEFINIRE LA NOZIONE DI PROBABILITA'

1. Introduzione

La teoria della probabilità costituisce uno strumento potente e nuovo per la formalizzare della realtà sia nell'ambito tradizionale delle scienze della natura sia in ambiti quali l'economia, la teoria dei giochi ed altro. Tutti abbiano una idea intuitiva di che cosa significhi che un evento sia più o meno probabile e questo è evidenziato dal fatto che continuamente prendiamo decisioni che tengono conto di tale intuizione. Facciamo valutazioni di tipo probabilistico sia se si tratta di portarci appresso l'ombrello quando usciamo di casa sia se si tratta di scegliere quali carte scartare in un giro di poker. Fortunatamente questa idea non rimane sempre a livello intuitivo e ben presto diventa un capitolo ben formalizzato della matematica che piano piano riguarderà non solo il gioco delle carte ma innumerevoli ed importati settori della ricerca scientifica.

Fissare una funzione di probabilità consiste nell' assegnare un numero ad un evento. Questo numero potrebbe essere considerato una misura della nostra propensione a ritenere che tale evento si verificherà (opinione di un soggetto). In questo caso entrano fortemente in gioco fortemente le informazioni che abbiamo a nostra disposizione. Ad esempio se dopo avere mescolato le carte di un mazzo di 40 carte mi viene data una carta coperta io posso valutare 4/40 = 1/10 la probabilità che l'asserzione "è un asso" sia vera. Se successivamente vengono scoperte 9 carte del mazzo e nessuna contiene un asso allora sono portato a ritenere molto maggiore tale probabilità, ad esempio 4/30. A limite se mi permettono di scoprire tutte le carte del mazzo allora posso dedurre esattamente se la carta coperta è un asso oppure no, e quindi valutare 0 o 1 la probabilità dell'evento "è un asso". Insomma la probabilità dipende dal mio livello di informazione.

Oppure una probabilità potrebbe essere considerata la misura di una intrinseca tendenza dell'evento a verificarsi, tendenza relativa alla natura del fenomeno fisico considerato e non alle conoscenze che ne abbiamo (grandezza fisica). In tale caso la probabilità dipende da "come va il mondo" e non dalle mie conoscenze del mondo. La meccanica quantistica utilizza la teoria della probabilità in tale senso.

Non è sorprendente pertanto che sulla natura della probabilità esistono, come vedremo, opinioni diverse. In ogni caso il punto di partenza di tutte le possibili definizioni è quello di “evento”. Si parte da un insieme Ω di possibili “eventi elementari”, ad esempio l’insieme delle sei facce di un dado o l’insieme delle carte da gioco di un mazzo. E’ possibile scommettere su un evento elementare, ad esempio nel caso dei dadi sull’evento “esce il numero 5”. Inoltre possiamo anche effettuare scommesse del tipo “esce un numero pari”, “esce un numero minore di 3” che corrispondono a scommettere sugli insiemi $\{2, 4, 6\}$ oppure $\{1,2\}$. Continuiamo a chiamare “eventi” gli insiemi di tale tipo e quindi accanto agli eventi elementari chiamiamo eventi anche i sottoinsiemi di Ω.

Come vedremo nel seguito, nel caso in cui Ω è infinito non tutti i sottoinsiemi possono essere considerati come eventi. Tuttavia si accetta che se di due eventi A e B è sensato parlare della probabilità che si verifichino allora è sensato parlare anche della probabilità che non si verifichino (negazione) o della probabilità che si verifichino entrambi (intersezione) o che si verifichi almeno uno di essi (unione). Ciò comporta che la classe degli eventi deve essere un algebra di Boole di insiemi.

Definizione 1.1. Dato un insieme Ω chiamiamo *algebra degli eventi* ogni sottoalgebra A dell’algebra di Boole $P(\Omega)$.

L’insieme Ω viene detto *evento certo*, l’insieme vuoto ∅ *evento impossibile*. I singoletti, se appartengono ad A, vengono detti *eventi elementari*. Due eventi X ed Y tali che $X \cap Y = \emptyset$ si dicono *incompatibili*.

La definizione ora data potrebbe risultare insufficiente. Infatti esiste l’esigenza di considerare il caso in cui Ω è l’insieme dei punti di uno spazio euclideo. Ciò comporta qualche complicazione a riferirsi a $P(\Omega)$ come insieme di tutti gli eventi. Infatti appare naturale accettare che la probabilità di colpire un punto di una figura F posta su un rettangolo R sia proporzionale alla misura di tale figura. Allora è ragionevole assumere come probabilità di colpire F il rapporto $m(F)/m(R)$ delle misure delle due figure. Tuttavia esistono sottoinsiemi F del piano euclideo che non sono misurabili (si veda il paradosso di Vitali esposto nel secondo volume) e pertanto in tali insiemi la probabilità non può essere definita. Da ciò segue che la classe degli insiemi misurabili non è chiusa per unioni qualsiasi poiché altrimenti essendo ogni insieme unione di punti, ogni insieme sarebbe misurabile.

Fortunatamente tale classe è chiusa per unioni finite o numerabili e questa proprietà è sufficiente ad applicare i potenti strumenti forniti dall'analisi matematica. In definitiva si mostra utile la seguente definizione.

Definizione 1.2. Dato un insieme Ω chiamiamo *σ-algebra degli eventi* ogni sottoalgebra *A* dell'algebra di Boole *P* chiusa per unioni numerabili.

2. Definizione classica (per i giochi d'azzardo)

Abbiamo già detto che matematizzare l'idea intuitiva di "probabile" consiste nell'assegnare ad ogni evento *A* un numero *p*(*A*). Esistono varie idee su come ciò possa essere fatto e sul significato da assegnare alla funzione *p*. Ne esponiamo le principali cominciando da quella che viene detta *definizione classica* o combinatoria.

Per introdurre tale definizione conviene ricordare che i primi studi matematici sulla probabilità sono legati ai giochi d'azzardo come quello dei dadi o quello delle carte. Essi sono presenti fin dalla prima metà del 1500 in *Liber de ludo aleœ* di Girolamo Cardano e se ne occupa anche Galileo Galilei nel suo *Sulla scoperta dei dadi* del 1656. Infatti Galileo elabora una spiegazione del perché lanciando tre dadi risultava empiricamente che la probabilità di ottenere un totale di 10 risultasse maggiore di quella di ottenere un totale di 9. Ciò a dispetto del fatto che sia il 9 che il 10 si ottengono con sei possibili combinazioni e precisamente

1+2+6, 1+3+5, 1+4+4, 2+2+5, 2+3+4), 3+3+3

e, rispettivamente

1+3+6, 1+4+5, 2+2+6, 2+3+5, 2+4+4, 3+3+4.

La cosa appariva paradossale poiché l'idea che si aveva della probabilità di un evento era che fosse uguale al rapporto tra il numero di casi in cui si verifica l'evento ed il numero totale di casi. Si applicava cioè la seguente definizione come verrà esplicitato da Laplace nei suoi lavori sulla probabilità.

Definizione 2.1. Sia l'insieme degli eventi uguale a $P(\Omega)$, allora la *funzione probabilità* $p : P(\Omega) \to [0,1]$ è definita ponendo

$$p(A) = |A|/|\Omega|.$$

In altre parole la probabilità di un evento è *il rapporto tra il numero dei casi favorevoli all'evento e il numero dei casi possibili*.

Ad esempio nel caso di un dado la probabilità che esca un numero pari è 3/6 e quindi 1/2. La probabilità che esca 5 è 1/6.

Dalla definizione seguono le seguenti proprietà.

Proposizione 2.2. Valgono le seguenti proprietà

1. la probabilità di un evento è un numero dell'intervallo [0,1],
2. la probabilità dell'evento certo è 1,
3. se A e B sono incompatibili allora $p(A \cup B) = p(A) \cup p(B)$.

Dim. La prima, seconda proprietà e quarta proprietà sono evidenti. Se $A \cap B = \varnothing$ allora

$$P(A \cup B) = |A \cup B|/|\Omega| = (|A|+|B|)/|\Omega|$$
$$= |A|/|\Omega|+|B|)/|\Omega| = p(A)+p(B). \quad \square$$

La proprietà 3 viene chiamata anche *finita additività*.

Tale definizione è perfetta e semplice da un punto di vista matematico. Purtroppo presenta notevoli aspetti negativi. Ad esempio:

- E' applicabile solo nei casi in cui tutti gli eventi si ritengono equiprobabili. Tale ipotesi potrebbe anche essere ragionevole nel caso dei giochi tuttavia la maggioranza dei fenomeni probabilistici non verifichino tale condizione. Per meglio dire è ragionevole nel caso di giochi non truccati. Ad esempio se un dado viene truccato in modo che la parte vicina al numero 6 sia più pesante della parte rimanente, allora è evidente che la probabilità che esca 6 è strettamente minore di 1/6 e non uguale ad 1/6.
- E' applicabile solo al caso in cui la classe degli eventi elementari è finita. Ad esempio esclude che si possa considerare la probabilità che un punto di un cerchio disegnato su un tabellone sia colpito se si spara a caso nella direzione del tabellone.

In ogni caso Galileo, tramite una corretta applicazione di tale definizione, chiarisce che per risolvere il paradosso dei tre dadi è sufficiente semplicemente riferirsi all'insieme corretto di eventi elementari.

Proposizione 2.3. (Soluzione del paradosso dei tre dadi). In un lancio di tre dadi la probabilità che la somma dei numeri usciti sia 9 è 25/216, la probabilità che esca 10 invece è 27/216.

Dim. Se i tre dadi sono distinti in primo, secondo e terzo dado, allora ogni volta che si gioca il risultato sarà rappresentato da una terna ordinata. Ad esempio la terna (1,4,4) significa che il primo dado ha fornito 1, il secondo 4 ed il terzo 4. Il numero di casi possibili è allora 6^3 che è il numero delle terne che si possono fare con sei elementi. La probabilità che esca 9 o 10 è data dalla percentuale di tali casi in cui ciò avviene, diviso ovviamente per 6^3. Ora per fare il calcolo corretto si deve tenere conto che:

- una combinazione con tre numeri uguali si presenta una volta sui 6^3 casi possibili,
- con due numeri uguali 3 volte
- con tre numeri diversi 6 volte.

Pertanto 9 si può ottenere in un numero di 6+6+3+3+6+1 = 25 casi mentre 10 si può ottenere in un numero di 6+6+3+6+3+3 = 27 casi. □

3. Definizione frequentista e definizione soggettiva

Una definizione diversa di probabilità è dovuta a Richard von Mises (1883-1953) e si basa sulla frequenza con cui un dato evento si verifica in una successione di esperimenti. Se n è il numero di volte in cui un esperimento si sia ripetuto (ad esempio il lancio di un dado) ed $f(A,n)$ il numero di volte che l'evento A si è verificato dopo n lanci, allora si considera la frequenza $f(A,n)/n$ come ipotesi sulla reale probabilità di A. Ad esempio se dopo avere lanciato il dado 100 volte mi accorgo che il numero 4 è uscito 42 volte, allora, in via provvisoria, ipotizzo che la probabilità di uscita di 4 è 42/100 e quindi che il dado è truccato non essendo gli eventi elementari equiprobabili. Von Mises riteneva inoltre che più a lungo veniva ripetuto l'esperimento più la frequenza calcolata era da considerare una misura attendibile della probabilità dell'evento A. Viene proposta pertanto la seguente definizione.

Definizione 3.1. (Definizione frequentista) La probabilità di un evento A è il limite cui tende la frequenza relativa di A al crescere del numero degli esperimenti:

$$p(A) = lim_{n\to\infty} f(A,n)/n.$$

Proposizione 3.2. La definizione frequentista di probabilità verifica le tre proprietà della proposizione 2.2.

Dim. Basta osservare che il limite di una successione costantemente uguale ad 1 è 1, che il limite di una successione di numeri compresi nell'intervallo [0,1] appartiene ancora a [0,1], e cheil limite della somma è la somma dei limiti. □

Tale definizione da un punto di vista matematico non ha molto senso.

- Infatti la successione $f(A,n)$ non è definibile da un punto di vista matematico ma è solo il risultato di una sequenza di esperimenti. Basta pensare che il valore di $f(A,7)$ risultante dall'esperimento di lanciare 7 volte il dado, è in genere diverso ogni volta che ripeto l'esperimento.
- Ammesso che la successione sia calcolabile non è detto che sia convergente.
- La nozione di "limite" delle frequenze relative non è paragonabile all'analogo concetto matematico poiché non esiste nemmeno dal punto di vista teorico la possibilità di calcolare il limite. Non è possibile nemmeno prendere una delle frequenze $f(A,p)/p$ come approssimazione del limite. Infatti questo è possibile solo se dato $\varepsilon > 0$ si sa trovare un numero naturale m tale che $|f(A,p)- f(A,q)| < \varepsilon$ per ogni $p > m$ e $q>m$.
- Tale definizione è applicabile solo ai fenomeni ripetibili (come quelli della fisica). Ma le valutazioni probabilistiche spesso si riferiscono a fenomeni non ripetibili. Ad esempio, ha sicuramente senso chiedersi quale sia la probabilità che l'attuale governo cada entro l'anno. Tuttavia non ha senso calcolare la probabilità di una sua caduta riferendoci ad una sequenza infinita di eventi uguali.

A dispetto di tali considerazioni la nozione di "frequenza" sembra appartenere comunque all'idea di probabilità. Forse si dovrebbe semplicemente reinterpretare la definizione 3.1 non come una definizione matematica ma come una descrizione di come in fisica si possa arrivare a proporre una funzione di probabilità in corrispondenza di un fenomeno casuale. In altre parole la descrizione di un "metodo" del tipo:

-se hai ripetuto l'esperimento un certo numero di volte allora è ragionevole assumere come misura di probabilità di un evento la frequenza con cui si è verificato.

Con la clausola successiva, che corrisponde al termine "limite",
- più volte hai fatto l'esperimento più è convincente tale assunzione.

La giustificazione che tale metodo funzioni, fortunatamente, non è dovuta quando si parla di fondamenti della matematica.

Passiamo ora ad una diversa definizione di probabilità che è stata proposta da De Finetti e Savage e che in genere viene chiamata *definizione soggettiva*.

Definizione 3.3. (Definizione soggettiva) La probabilità di un evento A è il prezzo $p(A)$ che un individuo ritiene equo pagare

- per ricevere 1 se l'evento si verifica,
- per ricevere 0 se l'evento non si verifica.

Tale prezzo deve essere "corretto" nel senso che non deve essere possibile una vincita o una perdita certa nel relativo sistema di scommesse.

Detto in questo modo sembra che le assegnazioni di probabilità siano qualcosa di completamente arbitrario. E' un errore poiché, come mostra il seguente teorema, l'ipotesi di "correttezza" comporta notevoli conseguenze.

Teorema 3.4. La definizione soggettiva di probabilità comporta la validità delle le proprietà 1, 2, 3 della proposizione 2.2.

Dim. Per provare che $p(A)$ è compresa tra 0 e 1 è sufficiente osservare che se $p(A)$ fosse negativa si avrebbe un guadagno certo, se fosse maggiore di 1 si avrebbe una perdita certa. Per provare che $p(\Omega) = 1$ osserviamo che se fosse $p(\Omega) < 1$ si avrebbe un guadagno certo pari a $1 - p(\Omega)$ poiché l'evento Ω si verifica sicuramente. Per provare infine la finita additività osserviamo per prima cosa che se gli eventi $A_1, A_2, ..., A_n$ costituiscono una partizione di Ω allora la somma delle probabilità $p(A_1)+\ldots+p(A_n)$ deve essere uguale ad 1. Infatti se fosse inferiore a 1 si avrebbe un guadagno certo, se fosse superiore si avrebbe una perdita certa. Consideriamo ora gli eventi incompatibili A e B, allora i tre insiemi A, B e $-(A\cup B)$ costituirebbero una partizione di Ω e quindi $p(A)+p(B)+p(-(A\cup B)) = 1$. Pertanto $p(A)+p(B) = 1- p(-(A\cup B))$. Poiché per quanto dimostrato sulle partizioni si ha che $p(A\cup B) = 1-p((A\cup B))$, possiamo concludere che $p(A\cup B) = p(A)+p(B)$. □

In accordo con tale teorema, possiamo dire che il punto di vista soggettivista non dice come definire la funzione p ma solo che tale funzione deve verificare alcune proprietà se vogliamo che il gioco sia ragionevole. Naturalmente essa non cade nei limiti della definizione classica e nemmeno di quella frequentista. Tuttavia pur essendo ragionevole nell'universo delle scommesse (e quindi, ad esempio, dell'economia) non è compatibile con l'idea per cui la probabilità è un fenomeno inerente molti fenomeni fisici. Questo è un limite non da poco se si pensa che in capitoli della fisica come la teoria dei gas e la meccanica quantistica la probabilità risulta indispensabile.

4. Definizione assiomatica

Non è sorprendente che in un periodo in cui emerge e domina una impostazione assiomatica della matematica, anche la nozione di probabilità venga assiomatizzata. La cosa viene fatta da Andrey Nikolaevich Kolmogorov nel 1933 ed in tale operazione si manifesta tutta l'efficienza del metodo assiomatico. Il sistema di assiomi che viene proposto coincide sostanzialmente con le tre proprietà che abbiamo fino ad ora visto, almeno nel caso finito.

Definizione 4.1. (Definizione assiomatica) Una probabilità *finitamente additiva* è una funzione $p : A \to [0,1]$ in una sottoalgebra A dell'algebra di Boole $P(\Omega)$ tale che $p(\Omega) = 1$ e che verifichi la finita additività, cioè

se $A\cap B = \varnothing$ allora $p(A\cup B) = p(A)\cup p(B)$.

In questo modo, come al solito, tramite il metodo assiomatico è possibile spostare l'attenzione da

"che cosa è la probabilità e come viene determinata"

a

"quali sono le proprietà (assiomi) che mi permettono di asserire che una funzione è una funzione di probabilità e quali sono i teoremi che posso dimostrare a partire da tali proprietà ?"

Abbiamo comunque osservato che per motivi tecnici quando si lavora nel continuo si deve supporre che l'algebra degli eventi sia una σ-algebra. In tale caso usualmente si accetta la seguente definizione che risulta quella più comunemente accettata oggi dai matematici.

Definizione 4.2. (Definizione assiomatica) Chiamiamo *probabilità numerabilmente additiva* ogni funzione $p : A \to [0,1]$ definita in una σ-algebra e tale che $p(\Omega) = 1$ e

$$p(\cup_{n\in N} A_n) = \sum_{n\in N} p(A_n)$$

per ogni successione $(A_n)_{n\in N}$ di elementi di A due a due disgiunti.

E' evidente che ogni probabilità numerabilmente additiva è anche finitamente additiva mentre si dimostra che non vale il viceversa.

5. Qualche nozione e qualche proposizione

Anche se questo libro non si pone certo lo scopo di fornire le nozioni base della teoria della probabilità, per potere continuare ad analizzare la nozione di probabilità abbiamo bisogno di fornire alcune semplici proposizioni e definizioni. Dagli assiomi per le probabilità è possibile ricavare alcune delle proprietà di una probabilità.

Proposizione 5.1. Se p è una probabilità finitamente additiva allora

a) $p(\emptyset) = 0$

b) $p(A)+p(-A) = 1$

c) se $A_1, A_2,..., A_n$ sono a due a due disgiunti allora

$$p(A_1\cup...\cup A_n) = p(A_1)+...p(A_n)$$

d) se $A\supseteq B$ allora $p(A-B) = p(A)-p(B)$

e) $p(A\cup B) = p(A)+p(B)-p(A\cap B)$

f) p è monotona

g) per ogni successione $(A_n)_{n\in N}$ di eventi a due a due disgiunti

$$p(\cup_{n\in N} A_n) \geq \sum_{n\in N} p(A_n) .$$

Dim. Per provare *a*) osserviamo che

$$p(\Omega) = p(\Omega\cup\emptyset) = p(\Omega)+p(\emptyset)$$

e quindi che $p(\emptyset) = p(\Omega)- p(\Omega) = 0$. Per provare *b*) osserviamo che

$$p(A)+p(-A) = p(A\cup -A) = p(\Omega) = 1.$$

La proposizione *c*) è evidente, per provare *d*) osserviamo che

$$P(A) = p((A-B)\cup B) = p(A-B)+p(B).$$

Per provare *e*) osserviamo che

$$p(A\cup B) = p(A\cup(B-(A\cap B)) = p(A)+p(B-A\cap B)$$
$$= p(A)+p(B)-p(A\cap B).$$

La proposizione *f*) segue da *e*). Infine per provare *g*) osserviamo che, poiché $\cup_{n\in N}A_n \supseteq A_1\cup ...\cup A_m$ per ogni m,

$$p(\cup_{n\in N}A_n) \geq p(A_1\cup ...\cup A_m) = p(A_1)+...p(A_m).$$ □

Supponiamo che io accetti una distribuzione di probabilità p ma che successivamente all'avere stabilito tale distribuzione io venga a conoscenza di nuove informazioni sull'evento elementare che si deve verificare. Ad esempio, adottando il punto di vista classico, inizialmente ritengo che nel gioco della tombola la probabilità che si realizzi l'evento $B = \{3, 78, 47\}$ sia di 3/90. Tuttavia, avendo osservato i numeri usciti, prendo atto che nel canestro è rimasto solo l'insieme A di numeri. Allora la probabilità che si verifichi l'evento $B = \{3, 78, 47\}$ subisce una variazione ed è ragionevolmente che sia posta uguale alla percentuale degli elementi di B che sono rimasti in A rispetto alla totalità degli elementi rimasti, cioè uguale a $|A\cap B|/|B|$. D'altra parte questo rapporo coincide con $p(A\cap B)/p(B)$. Infatti

$$p(A\cap B)/p(B) = (|A\cap B|/90)\cdot(90/|B|) = |A\cap B|/|B|.$$

Ciò suggerisce la seguente definizione.

Definizione 5.2. Data una probabilità finitamente additiva p ed un evento B tale che $p(B) \neq 0$, chiamiamo *probabilità di A condizionata da B* il numero

$$p(A|B) = p(A\cap B)/p(B).$$

Si osservi che la probabilità di A condizionata da B può essere sia maggiore che minore della probabilità primitiva $p(A)$ assegnata ad A. Ad esempio se A e B si escludono a vicenda, cioè se $A\cap B = \emptyset$, allora $p(A|B) = 0$. Se $B\subseteq A$ allora $p(A|B) = 1$.

Una ulteriore nozione fondamentale per la teoria della probabilità è quella di valore di aspettazione o valore medio di una grandezza. La questione è la seguente. Supponiamo che nel gioco dei dadi si stabilisca che ogni volta che esce un numero n si riceva una posta di n^2 e che per partecipare al gioco si paghi una quota r. Allora si pone il problema di quanto debba essere tale quota perché il tutto sia corretto. Ora se il dado non è truccato ci si aspetta che in un lancio ripetuto si ottenga 1/6 delle volte 1 (guadagno 1), un sesto delle volte 2 (guadagno 2^2), ... 1/6 delle volte 6 (guadagno 6^2). In totale ci si aspetta che ci sia un guadagno di

$$1/6 \cdot 1^2 + 1/6 . 2^2+1/6 \cdot 3^2 + 1/6 . 4^2 + 1/6 . 5^2+1/6 . 6^2$$

pari a 91/6. Quindi la quota corretta da pagare dovrebbe essere pari a 91/6. Ebbene, tale numero rappresenta il valore medio o valore di aspettazione della funzione che ad ogni evento elementare associa la corrispondente vincita. Si arriva pertanto alla seguente definizione.

Definizione 5.3. Sia Ω finito, p una probabilità finitamente additiva in $P(\Omega)$ ed $f : \Omega\rightarrow R$ una funzione a valori reali. Chiamiamo *valore di aspettazione* o *valore medio* di f il numero

$$\sum_{x\in\Omega} p(x)\cdot f(x).$$

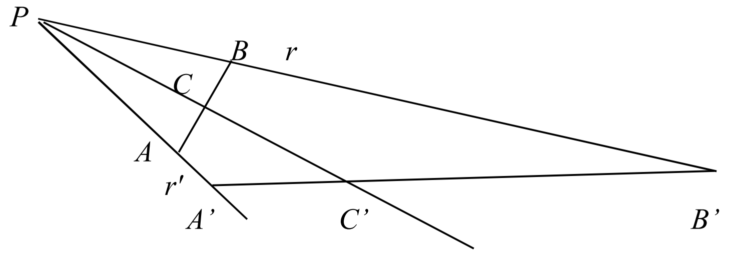
Nel caso in cui l'insieme Ω non sia finito si deve ricorrrere all'analisi matematica sostituendo alla sommatoria un opportuno integrale della funzione f.

6. Alcuni paradossi della probabilità

Espongo alcuni paradossi senza tentare di trovarne una soluzione con la speranza che chi legge si sforzi di trovarne una. Lo scopo è solo di mettere in evidenza quanto sia problematica la nozione di probabilità.

Paradosso dell'equipotenza. Da un punto P si traccino due rette r ed r' e poi la bisettrice dell'angolo che si è formato. Costruiamo una perpendicolare a tale bisettrice ed indichiamo con A, C e B i tre punti di intersezione di tale perpendicolare con le tre rette per P. Per costruzione AC è uguale a CB. Si tracci inoltre una retta che intercetti tre punti A', C' e B' in modo che $A'C'$ sia molto minore di $C'B'$. Detto questo supponiamo che dal punto P vengano sparati a caso dei colpi di pistola in tutte le direzioni ma sempre all'interno dell'angolo tra le rette r ed r'. E' evidente che la probabilità i colpire $A'C'$ è molto minore della probabilità di colpire $C'B'$ essendo il primo segmento di minore estensione.

Inoltre la probabilità di colpire AC è la stessa della probabilità di colpire CB essendo i due segmenti della stessa estensione.



Abbiamo quindi che $p(AC) = p(CB)$ e $p(A'C') < p(C'B')$. D'altra parte tutte le volte che un proiettile colpisce AC colpisce necessariamente anche $A'C'$ e tutte le volte che colpisce AC' deve necessariamente avere colpito AC. Pertanto $p(A'C') = p(AC)$ e per lo stesso motivo $p(C'B') = p(CB)$. Da ciò segue che $p(C'B') = p(A'C')$ in contrasto con il fatto che $p(A'C') < p(C'B')$.

Ho chiamato "paradosso dell'equipotenza" questo paradosso poiché si basa sull'idea che due eventi A e B che sono equipotenti dovrebbero avere la stessa probabilità di verificarsi. Infatti sia $f: A \to B$ è una funzione biettiva tra A e B e supponiamo che esista un marchingegno che ogni volta che si verifica un evento elementare x crei l' evento elementare $f(x)$. Allora accade che la probabilità che si verifichi A non può che essere uguale alla probabilità che si realizzi B. Questo fatto è compatibile con la concezione classica della probabilità in cui l'ipotesi di equiprobabilità per gli eventi elementari comporta l'equiprobabilità tra eventi equipotenti. Porta invece a contraddizione se si definisce la probabilità come rapporto di misure nello spazio euclideo. Infatti due insiemi possono essere equipotenti ma con misure diverse.

Paradosso degli insiemi di misura nulla. Sia F un insieme numerabile di punti di un quadrato Q. Allora essendo F di misura nulla la probabilità di colpire un punto di F sparando a caso verso il quadrato è zero. Inoltre, poiché Q-F ha misura 1, siamo certi che tutti i colpi cadranno in Q-F. Questo tipo di paradosso nasce dal fatto che la nozione di area, o più in generale di misura in un insieme infinito, non è abbastanza "fine" da assicurare che un insieme di misura nulla sia necessariamente l'insieme vuoto.

Paradosso di San Pietroburgo

Questo paradosso è stato esposto dal matematico Daniele Bernoulli e riguarda un gioco basato sul lancio di una moneta. Si lancia una moneta fino a quando non esce Testa. La vincita dipende dal numero di lanci necessari: se esce *Croce* al primo lancio, si vince 1 (uno), se esce all'n-esimo lancio si vince 2^n. Se non esce mai ovviamente non si vince. Quale è la cifra corretta che si dovrebbe pagare per partecipare a tale gioco ? Tale cifra dovrebbe coincidere con la vincita media che ci si aspetta di avere, cioè con il valore di aspettazione della funzione-vincita. I casi possibili sono i seguenti.

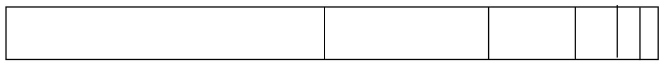
- esce testa al primo lancio: probabilità 1/2, guadagno 1
- esce testa al secondo lancio: probabilità $(1/2)^2$, guadagno 2
- . . .
- esce testa al lancio n: probabilità $(1/2)^n$, guadagno n
- ...
- non esce mai testa, guadagno 0

pertanto il valore di aspettazione è

$$V = (1/2)\cdot 1+(1/2)^2\cdot 2+(1/2)^3\cdot 2^2+\ldots.+(1/2)^n\cdot 2^{n-1}+\ldots\ldots$$
$$= 1/2+1/2+\ldots = \infty$$

Pertanto qualunque cifra venga pagata dal giocatore per partecipare a questo gioco si dimostra inadeguata. Ciò è fortemente contrario all'intuizione che invece suggerisce che tale gioco non sia poi "infinitamente conveniente".

Possiamo anche dare una forma "geometrica" a tale paradosso. Supponiamo di sparare a caso su di una tavola di legno



lunga un metro e che la tavola sia stata divisa in rettangoli A_1, A_2, ... come in figura e che ogni rettangolo sia la metà del precedente. Il gioco consiste nel fatto che viene sparato a caso un colpo verso la tavola. Se viene colpito il rettangolo A_n allora si riceve la cifra n. Se si tiene conto che la probabilità di colpire la tavola A_n è $(1/2)^n$, esattamente con gli stessi calcoli che abbiamo fatto per il paradosso di San Pietroburgo, abbiamo che il guadagno medio che ci si aspetta di avere è infinito. Quindi qualunque quota si voglia fissare per partecipare a questo gioco si dimostra inadeguata.

Paradosso delle due buste
Questo paradosso è stato ed è oggetto di numerose discussioni tra i probabilisti. Si supponga di avere due buste chiuse uguali e che in una sia stata inserita una cifra e nell'altra il doppio di tale cifra. Non si conoscono le cifre messe in gioco. Il giocatore deve scegliere una busta, aprirla e decidere se preferisce tenersi la somma che ha trovato oppure prendere la somma, sconosciuta, che è nella busta rimasta chiusa. Quale è la scelta più razionale ? Supponiamo ad esempio che nella busta aperta trova la cifra di 1000 euro. Allora la scelta della busta rimasta chiusa comporta un valore medio di guadagno pari a $V = 1/2{\cdot}2000+1/2{\cdot}500 = 1/2{\cdot}2500 = 1250$. La scelta della busta già aperta un valore di 1000. Quindi risulta sempre conveniente aprire la busta rimasta chiusa. Ciò è molto strano poiché le due buste non sono distinguibili e non si capisce perché la seconda sia preferibile alla prima.

Paradosso delle tre carte colorate
Ci sono tre carte: una ha ambedue le facce rosse, una ha una faccia rossa e una nera, la terza ha ambedue le facce nere. Le metto in un sacchetto, le mischio e, a occhi chiusi, ne tiro fuori una e l'appoggio sul tavolo. Apro gli occhi e mi accorgo che la faccia visibile è rossa. Qual'è la probabilità che anche l'altra faccia sia rossa? Esistono due modi alternativi di procedere

Primo tipo di calcolo - Io vedo una faccia rossa per cui escludo la carta nera/nera. Rimangono due alternative: o ho estratto la carta rossa/nera oppure quella rossa/rossa, ugualmente probabili. Pertanto la probabilità è del 50%.

Secondo tipo di calcolo - La faccia rossa che vedo può essere una delle tre facce rosse esistenti:
o una delle due facce rosse della carta rossa/rossa,
o l'altra faccia (sempre rossa) della stessa carta
oppure l'unica faccia rossa della carta rossa/nera.
Solo in quest'ultimo caso (su tre) dall'altro lato c'è una faccia nera. Pertanto la probabilità che l'altra faccia sia rossa è 2/3 cioè 66%.

Avviene pertanto che due metodi diversi di calcolo della probabilità forniscono risultati diversi.

Lo stesso problema può essere riformulato al modo seguente. Prendo a caso una delle tre carte e la pongo su di un tavolo coprendola con una mano. Poi chiamo un amico e gli dico: se la carta estratta ha due facce dello stesso colore hai vinto un milione altrimenti non hai vinto niente. Il mio amico calcola che poiché le carte con lo stesso colore sono 2 su 3, la probabilità di vincere un milione è 2/3. Detto questo dico che per prima cosa alzerò la mano mostrando uno dei lati della casta, poi scoprirò la carta per vedere se è una carta con lo stesso colore sui due lati. A questo punto il mio amico fa una richiesta strana: "preferirei che tu mi dicessi in una sola volta se è una carta giusta oppure no". Il ragionamento che ha fatto è il seguente. Se la carta poggiata sul tavola è ross allora il caso nero/nero è escluso e quindi restano solo le possibilità rosso/rosso e rosso/nero. Pertanto la probabilità passa da 2/3 ad 1/2 . La stessa cosa avviene ovviamente se la carta è nera. Questo fatto ovviamente è paradossale poiché non si vede perchè due procedure diverse per vedere la carta possano influenzare la probabilità di vincita del mio amico.

Paradosso della conferma
Questo paradosso è stato ideato dal logico tedesco C. G. Hempel nel 1937. In realtà non è un paradosso della teoria della probabilità, piuttosto è un paradosso logico il quale (forse) può essere risolto grazie alla teoria della probabilità.

Consideriamo l'affermazione "Tutti i corvi sono neri" che possiamo riscrivere in termini logici con

$$\text{“}\forall x(x \text{ è un corvo} \Rightarrow x \text{ è nero)”.} \qquad (*)$$

In quanto affermazione scientifica su dati di fatto, una sua conferma si ottiene tutte le volte che vediamo un corvo e constatiamo che è nero. Ora in logica matematica sappiamo che una implicazione $C \Rightarrow N$ è equivalente alla sua "contronominale" $\neg N \Rightarrow \neg C$ nel senso che la prima è vera se e solo se la seconda è vera. Quindi l'implicazione (*) equivale a

$$\text{“}\forall x(x \text{ non è nero} \Rightarrow x \text{ non è un corvo)”.} \qquad (**)$$

Ma questo comporta che la legge (*) è confermata tutte le volte che vedo un animale non nero e mi accorgo che non è un corvo. Ad esempio tutte le volte che vedo un coniglio bianco !

In realtà la stranezza non ci sarebbe se ci riferissimo ad insiemi piccoli di animali ad esempio gli animali di una isoletta.[1] Ma quando enunciamo una regola del tipo "tutti i corvi sono neri" intendiamo enunciare una regola universale che vale non solo per gli animali attualmente esistenti ma per tutti gli animali che possono esistere in futuro. Inoltre è evidente che la regola (*) non viene inficiata nel caso si trovi un solo corvo "albino", cioè nel caso che per qualche caso strano sia nato un animale che per tutti gli aspetti è da ritenere un corvo ma che risulta non nero.

Un modo per risolvere il paradosso della conferma è quello di interpretare l'affermazione "tutti i corvi sono neri" in termini probabilistici cioè come "è probabile che se mi appare un corvo allora tale corvo risulti nero". Se indico con C l'evento "essere un corvo" e con N l'evento "essere nero" allora la misura di tale probabilità è rappresentata dalla probabilità condizionata $p(N|C)$. In tale modo il paradosso si blocca poiché l'equivalenza tra una implicazione e la sua contronominale non sussiste nel caso delle probabilità condizionate. In altre parole, non vale l'uguaglianza $p(N|C) = p(-C|-N)$. Ad esempio nell'insieme delle vocali posso porre $C = \{a,e,o\}$ e $N = \{e,i,o\}$, allora $p(N|C) = |\{e,o\}|/|A| = 2/3$ mentre $p(-C|-N) = |\{u\}|/|\{a,u\}| = 1/2$.

Paradosso di Carlo V.

Leggendo la Settimana Enigmistica ho trovato la seguente "Spigolatura".

Nel 1952, durante la campagna militare che oppose Enrico II a Carlo V, quest'ultimo stava ispezionando la propria armata, insediata davanti a Metz, quando si avvicinò un po'troppo ad una delle batterie nemiche. Numerosi ufficiali del suo esercito lo pregarono allora di spostarsi in una posizione più sicura, ma Carlo V esclamò ridendo: «rassicuratevi: potete citarmi un solo imperatore che sia stato colpito da una palla di cannone?».

Tale spigolatura determina un paradosso che ho pensato di chiamare *"paradosso di Carlo V"*. Infatti il numero che misura la probabilità di un evento viene usualmente assunto uguale alla

---

[1] Anzi, se esistono 20 corvi e 5 animali bianchi che non sono corvi, allora per verificare (*) ho bisogno di 20 verifiche, per verificare (**) ho bisogno di 5 verifiche. Pertanto (**) risulta non solo equivalente ma anche "più conveniente" di (*).

percentuale dei casi passati in cui si è verificato l'evento. Pertanto il comportamento spericolato di Carlo V è in completo accordo con la teoria della probabilità. D'altra parte con tale tipo di argomentazione Carlo V potrebbe posizionarsi ad un metro dal cannone senza correre alcuni rischio e questo è assurdo.

Paradosso delle due bombe
Carlo, famoso studioso di teoria della probabilità, viaggia spesso per i suoi convegni ma ha sempre una grande paura che un terrorista faccia esplodere l'aereo in cui si imbarca. Un giorno un poliziotto della sorveglianza lo perquisisce e gli trova un bomba nella borsa…

- Lei è in arresto, perché porta questa bomba ?
- E Lei che è un ignorante ! Ho fatto i miei calcoli e risulta che la probabilità che ci sia una bomba in un aereo è di uno su un milione ma la probabilità che ci siano due bombe è di uno su un miliardo. Quindi se io porto una bomba nel mio aereo sono quasi sicuro che non ce ne sono altre !

APPENDICE
NOZIONI BASE E VARIE

1. Coppie, prodotti cartesiani e relazioni

Questo ed il prossimo volume sono rivolti a persone che hanno già una conoscenza elementare della matematica. Tuttavia, per permetterne una lettura ad una platea la più ampia possibile, in questa appendice ricordo alcune delle nozioni utilizzate. Nel seguito indicherò:

- con $\emptyset$ l'insieme vuoto,
- con $\{d_1,\dots,d_n\}$ l'insieme i cui elementi sono $d_1,\dots,d_n$.
- con $\{x : x \text{ verifica la proprietà } P\}$ l'insieme i cui elementi sono tutti e soli quelli verificanti la proprietà P.

Il primo passo per la definizione dei concetti fondamentali della teoria degli insiemi è quello di definire la nozione di coppia ordinata e di prodotto cartesiano.

Definizione 1.1. Dati due elementi x ed y chiamiamo *coppia di primo elemento x e secondo elemento y* l'insieme $\{\{x\},\{x,y\}\}$ che indichiamo con (x,y). Dati due insiemi X ed Y chiamiamo *prodotto cartesiano* l'insieme $X\times Y = \{(x,y) : x\in X, y\in Y\}$ delle coppie costituite da un elemento di X ed un elemento di Y.

Da notare che gli insiemi $\{\{x\},\{x,y\}\}$, $\{\{x,y\},\{x\}\}$ $\{\{y,x\},\{x\}\}$ e $\{\{x\}, \{x,y\}\}$ coincidono tra loro e rappresentano tutti la stessa coppia (x,y). Inoltre la coppia (x,x) è rappresentata dall'insieme $\{\{x\}\}$.

Definizione 1.2. Chiamiamo relazione *binaria* tra un insieme X ed un insieme Y ogni sottoinsieme $\mathcal{R}$ di $X\times Y$. L'*inversa* di $\mathcal{R}$ è la relazione $\mathcal{R}^{-1} = \{(x,y) : (y,x)\in \mathcal{R}\}$. Date due relazioni $\mathcal{R}_1\subseteq X\times Y$ e $\mathcal{R}_2\subseteq Y\times Z$, la loro *composizione* è la relazione $\mathcal{R}_1\circ\mathcal{R}_2$ definita ponendo

$$\mathcal{R}_1\circ\mathcal{R}_2=\{(x,z)\in X\times Z : \text{esiste } y\in Y,\ (x,y)\in\mathcal{R}_1,(y,z)\in\mathcal{R}_2\}.$$

Definizione 1.3. Data una relazione $\mathcal{R}$ il suo *dominio* viene definito ponendo

$$Dom(\mathcal{R}) = \{x\in X : \text{esiste } y\in Y \text{ per cui } (x,y)\in\mathcal{R}\},$$

il suo *codominio* ponendo

$Cod(\mathcal{R}) = \{y \in Y \ : \ x \in X$ per cui $(x,y) \in \mathcal{R}\}$.

Usualmente per le relazioni binarie si utilizza la *notazione infissa* che consiste nello scrivere $x\mathcal{R}y$ per indicare che $(x,y) \in \mathcal{R}$.

Definizione 1.4. Una relazione binaria $f \subseteq X \times Y$ è chiamata *funzione di X in Y* se è univoca, cioè se, per ogni $x \in X$ esiste al più un elemento y tale che $(x,y) \in f$. Se per ogni $x \in X$ esiste uno ed un solo elemento y tale che $(x,y) \in f$, cioè se $Dom(f) = X$ allora diciamo che f è *totale*, altrimenti che è *parziale*.

In analisi matematica $Dom(f)$ viene chiamato anche *campo di esistenza* di f. Nel seguito denoteremo una funzione con una lettera minuscola, ad esempio la lettera f e scriveremo $f : X \rightarrow Y$ per indicare che f è una funzione di X in Y. Per ogni $x \in X$ indichiamo con $f(x)$ l'unico elemento y tale che $(x,y) \in f$. In alcuni testi l'insieme X viene detto *dominio*, l'insieme Y viene detto *codominio* di f. In altri testi per dominio e codominio si intendono i due insiemi $Dom(f)$ e $Cod(f)$.

Definizione 1.5. Una funzione $f : X \rightarrow Y$ è *iniettiva* se

$$f(x) = f(y) \Rightarrow x = y,$$

è *suriettiva* se per ogni $y \in Y$ esiste almeno un $x \in X$ tale che $f(x) = y$ cioè se $Cod(f) = Y$, è *biettiva* se è sia suriettiva che iniettiva.

2. Definizione (brutta) di *n*-pla

Anche se in tutti i libri viene proposta la definizione di coppia che abbiamo dato nel paragrafo precedente, ci si pone il problema di perché si sia optato per una definizione così strana, arbitraria e comunque poco intuitiva. Ancora meno intuitiva è la nozione di n-pla che viene fatta per induzione su $n \geq 2$ al modo seguente.

Definizione 2.1. Dato $n \geq 2$ e gli insiemi $X_1,...,X_n$ definiamo il *prodotto cartesiano* di tali n insiemi per ricorsione su n tramite l' equazione

$$X_1 \times ... \times X_n = (X_1 \times ... \times X_{n-1}) \times X_n$$

e chiamando *n-pla* un elemento di tale prodotto cartesiano. Se tutti gli insiemi $X_1,...,X_n$ coincidono con l'insieme X allora denotiamo con X^n tale prodotto cartesiano.

Pertanto una terna $(x,y,z) = ((x,y),z)$ viene a coincidere con l'insieme

$$\{\{\{x\},\{x,y\}\},\{\{\{x\},\{x,y\}\},z\}\}.$$

In particolare la terna (x,x,x) coincide con l'insieme $\{\{\{x\}\},\{\{\{x\}\},x\}\}$. Lascio a chi legge il divertimento di dire che cosa è una quadrupla (x,x,x,x). Solo i matematici sono capaci di rappresentare una cosa tanto semplice in modo tanto tortuoso ! La questione è che i matematici hanno introdotto la teoria degli insiemi per il desiderio di trovare uno strumento unico per costruire tutta la matematica. Quindi sono costretti a rappresentare quella che è intuitivamente una coppia come (2,5) usando solo strumenti insiemistici e quindi solo le parentesi {,} ed i simboli 2,5. La rappresentazione deve essere tale da rendere possibile la "estrazione" dell'informazione relativa a quale si intenda come primo e quale come secondo elemento. La Definizione 2.1 permette appunto di ottenere questo. La nozione di n-pla per $n>2$ può essere invece semplificata ricorrendo a quella di funzione. Ad esempio una terna potrebbe essere definita come una funzione definita in $\{1,2,3\}$ dopo avere indicato con 1, 2, 3 gli insiemi $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}$. D'altra parte quando pensiamo ad una n-pla $(x_1,...,x_n)$ pensiamo appunto ad una corrispondenza che associa ad ogni "posto" in $\{1,...,n\}$ un elemento. Se $X_1,...,X_n$ sono insiemi allora il prodotto cartesiano verrebbe definito ponendo

$$X_1\times...\times X_n = \{f: \{1,...,n\}\rightarrow X_1\cup...\cup X_n \text{ tali che } f(i)\in X_i\}.$$

Tale modo di definire il prodotto cartesiano si presta bene ad essere generalizzata al caso infinito al modo seguente.

Definizione 2.2. Sia $(S_i)_{i\in I}$ una famiglia di insiemi. L'insieme delle applicazioni $f: I\rightarrow \cup_{i\in I}S_i$ tali che $f(i)\in S_i$, per ogni $i\in I$, viene detto *prodotto cartesiano* della famiglia $(S_i)_{i\in I}$ e lo si indica con $\times_{i\in I}S_i$.

Fortunatamente i matematici dopo avere imparato la definizione insiemistica di n-pla la dimenticano rapidamente per tornare a basarsi, nei propri ragionamenti, sulla nozione intuitiva che tutti abbiamo.

La nozione di prodotto cartesiano permette di definire le nozioni generali di relazione n-aria e di operazione n-aria.

Definizione 2.3. Siano $X_1,\ldots,X_n$ insiemi non vuoti. Allora chiamiamo *relazione n-aria* ogni sottoinsieme del prodotto cartesiano $X_1\times\ldots\times X_n$. Dato un insieme non vuoto X chiamiamo *relazione n-aria in X* ogni sottoinsieme di X^n. Chiamiamo *operazione n-aria* in X una funzione di X^n in X.

3. Relazioni di equivalenza e quozienti

Esistono proprietà particolarmente importanti per le relazioni binarie su di un dato insieme. Ne elenchiamo alcune:

Definizione 3.1. Dato un insieme S ed una relazione binaria $\mathcal{R}$ in S diremo che

- $\mathcal{R}$ è *riflessiva* se $x\mathcal{R}x$ per ogni $x\in S$
- $\mathcal{R}$ è *transitiva* se per ogni $x,y,z\in S$
  $x\mathcal{R}y$ e $y\mathcal{R}z \Rightarrow x\mathcal{R}z$
- $\mathcal{R}$ è *simmetrica* se per ogni $x,y\in S$
  $x\mathcal{R}y \Rightarrow y\mathcal{R}x$
- $\mathcal{R}$ è *asimmetrica* se per ogni $x,y\in S$
  $x\mathcal{R}y$ e $y\mathcal{R}x \Rightarrow x = y$
- $\mathcal{R}$ è *totale* o *lineare* se per ogni $x, y\in S$
  $x\mathcal{R}y$ oppure $y\mathcal{R}x$.

Definizione 3.2. Una relazione binaria $\mathcal{R}$ in un insieme S è una *relazione di equivalenza* se è riflessiva, transitiva e simmetrica.

In generale indicheremo una relazione di equivalenza con $\equiv$.

Esercizio. Dato un insieme non vuoto S, dimostrare che l'identità, cioè la relazione

$$D(S) = \{(x,y) : x \text{ ed } y \text{ sono lo stesso elemento di } S\}$$

è una relazione di equivalenza (la più piccola su S).

Esercizio. Dimostrare che la relazione totale, cioè la relazione $\mathcal{R} = S\times S$ secondo cui tutti gli elementi sono equivalenti tra loro è una relazione di equivalenza (la più grande su S).

Le relazioni di equivalenza sono alla base delle *definizioni per astrazione* in cui si considerano come un unico oggetto oggetti che, pur essendo diversi, differiscono per aspetti che si considerano non essenziali. Ad esempio consideriamo il seguente enunciato di un problema:

calcolare l'area di un triangolo di lati 3, 4, 5.

Spesso lo stesso problema viene enunciato al modo seguente:

calcolare l'area del triangolo di lati 3, 4, 5.

Questa seconda formulazione del problema è corretta ? Il fatto che esistono infiniti triangoli con lati 3,4,5 sembra mostrare che non lo sia. Tuttavia l'uso del singolare significa che si è deciso di considerare un nuovo oggetto "astratto" *il* triangolo di lati 3, 4, 5 che in un certo senso rappresenta tutti i possibili triangoli di lati 3, 4, 5. Ciò è possibile poiché, per quanto riguarda il problema da affrontare, non è interessante sapere la posizione di un triangolo sul piano ma solo le sue dimensioni. Alla base di tale processo è la relazione di equivalenza per cui due triangoli sono *uguali* se hanno lati uguali. Per le strutture algebriche si procede in modo analogo. E possibile dire un gruppo di ordine 5 ma è possibile dire anche il gruppo di ordine 5. Infatti tutti i gruppi di ordine 5 sono isomorfi tra loro e l'isomorfismo è una relazione di equivalenza.

Un modo per formalizzare un tale modo di procedere è identificare un oggetto astratto definito in questo modo con l'insieme degli oggetti *concreti* da esso rappresentato.

Definizione 3.3. Data una relazione di equivalenza $(S, \equiv)$ ed $x \in S$, la *classe completa di equivalenza determinata* da x è definita ponendo

$$[x] = \{x \in S : x' \equiv x\}.$$

Il *quoziente di S modulo* $\equiv$ è l'insieme $S/\equiv$ delle classi complete di equivalenza, cioè

$$S/\equiv\ = \{[x] : x \in S\}.$$

In altre parole se parto da un universo di oggetti S ed introduco una relazione di equivalenza $\equiv$ tra oggetti di tale insieme, allora vengo a creare per astrazione un nuovo insieme di oggetti $S/\equiv$. Tornando all'esempio dei triangoli, l'espressione *"calcolare l'area del triangolo di lati* 3, 4, 5*"* diventa corretta se col termine "triangolo di lati 3, 4, 5" si intende un unico oggetto: la classe completa di equivalenza costituita da tutti i triangoli i cui lati hanno tali lunghezze.

Proposizione 3.4. Siano S ed S' due insiemi ed $f : S \to S'$ una funzione. Allora la relazione $\equiv$ definita ponendo

$$x \equiv y \Leftrightarrow f(x) = f(y)$$

è una relazione di equivalenza che viene detta il *nucleo* di f.

Dim. La dimostrazione si lascia per esercizio. □

In altri termini, data una funzione f, il nucleo è la relazione "avere la stessa immagine". Nell'esempio dei triangoli la relazione di eguaglianza è determinata dalla funzione che associa ad ogni triangolo la terna costituita dalla lunghezza dei suoi lati. D'altra parte questa è l'esatta trascrizione della usuale definizione "due triangoli si dicono uguali se hanno lati uguali".

Esempio: Sia S l'insieme i cui elementi sono mucchietti di monete e sia f la funzione che associa ad ogni mucchietto x la somma totale rappresentata da x. Allora due mucchietti sono da considerare equivalenti se corrispondono allo stesso valore.

La nozione di nucleo è di notevole importanza e caratterizza le relazioni di equivalenza. Infatti vale la seguente proposizione.

Proposizione 3.5. Ogni relazione di equivalenza è il nucleo di una opportuna funzione.

Dim. Sia $\equiv$ una qualunque relazione di equivalenza in un insieme S e sia $S' = S/\equiv$. Allora la funzione $f : S \to S'$ ottenuta ponendo $f(x) = [x]$ è tale che

$$f(x) = f(y) \Leftrightarrow [x] = [y] \Leftrightarrow x \equiv y.$$

Ciò mostra che $\equiv$ è il nucleo di f. □

4. Relazioni d'ordine, reticoli ed algebre di Boole

Una classe importante di relazioni binarie in un insieme sono le relazioni d'ordine.

Definizione 4.1. Diciamo che una relazione binaria $\mathcal{R}$ in un insieme S è una *relazione di preordine* se è riflessiva e transitiva. Diciamo che $\mathcal{R}$ è una *relazione d'ordine* se è riflessiva, transitiva ed asimmetrica.

Le relazioni di preordine e di ordine usualmente vengono denotate con il simbolo $\leq$. Se S è un insieme ed $\mathcal{R}\subseteq S\times S$ una relazione binaria su S allora a volte si usa dire che la coppia $(S,\mathcal{R})$ è una *struttura relazionale*. In accordo con la definizione generale di immersione e di isomorfismo tra strutture relazionali che esporremo nell'ultimo paragrafo, abbiamo la seguente definizione.

Definizione 4.2. Date due strutture relazionali $(S_1,\mathcal{R}_1)$ e $(S_2,\mathcal{R}_2)$ chiamiamo *immersione* di $(S_1,\mathcal{R}_1)$ in $(S_2,\mathcal{R}_2)$ ogni funzione biettiva $f : S_1\rightarrow S_2$ tale che

$$(x,y)\in\mathcal{R}_1 \Leftrightarrow (f(x),f(y))\in\mathcal{R}_2.$$

Diciamo che f è un *isomorfismo* se è una immersione suriettiva.

Definizione 4.3. Data una struttura di preordine $(S,\leq)$ ed un sottoinsieme X di S prende il nome di *maggiorante* (*minorante*) di X un elemento $m\in S$ tale che $x\leq m$ ($m\leq x$) per ogni $x\in X$. Prende il nome di *massimo* (*minimo*) *elemento* di X un elemento $m\in X$ tale che $m\geq x$ ($x\geq m$) per ogni $x\in X$. Si chiama *estremo superiore* un elemento $sup(X)$ che sia il minimo dell'insieme dei maggioranti di X. Si chiama *estremo inferiore* un elemento $inf(X)$ che sia il massimo dell'insieme dei minoranti di X.

Non è detto che gli estremi superiori o inferiori esistano sempre. Ad esempio l'insieme dei numeri primi non ammette estremo superiore.[1] Se m è il minimo di X è anche l'estremo inferiore ma il viceversa non vale. La stessa cosa si può dire per il massimo. Ad esempio se X è l'intervallo aperto (0,1) allora 0 è l'estremo inferiore ma non è il minimo mentre 1 è l'estremo superiore ma non è il massimo.

Definizione 4.4. (Definizione relazionale) Si chiama *reticolo* un insieme ordinato $(S,\leq)$ tale che per ogni coppia x ed y di elementi di S esistono $sup\{x,y\}$ e $inf\{x,y\}$. Si dice che $(S, \leq)$ è *limitato* se esistono in S il minimo e massimo che denotiamo con 0 ed 1, rispettivamente. Un reticolo è *completo* se per ogni insieme X di elementi di S esistono $sup(X)$ e $inf(X)$.

[1] Ricordiamo che gli antichi greci provarono che per ogni numero primo p esiste un numero primo q maggiore di p (in termini attuali diremmo che l'insieme dei numeri primi è infinito).

Dato un reticolo completo esiste l'estremo superiore della famiglia di tutti gli elementi di S. Tale estremo superiore è ovviamente il massimo di S ed usualmente viene denotato con 1. Similmente esiste l'estremo inferiore della famiglia di tutti gli elementi di S. Tale estremo inferiore è il minimo di S e viene usualmente denotato con 0.

Spesso è più comodo definire un reticolo come struttura algebrica cioè come insieme più due operazioni.

Definizione 4.5. (Definizione algebrica) Chiamiamo *reticolo* una struttura algebrica $(S,\vee,\wedge)$ che verifica i seguenti assiomi

(i) $x\vee y = y\vee x$; $x\wedge y = y\wedge x$ (commutativa)
(ii) $x\vee(y\vee z) = (x\vee y)\vee z$; $x\wedge(y\wedge z)=(x\wedge y)\wedge z$ (associativa)
(iii) $x\vee x = x$; $x\wedge x = x$ (idempotenza)

Le due diverse definizioni di reticolo non creano confusione poiché valgono le due seguenti proposizioni.

Proposizione 4.6. Ad ogni struttura relazionale $(S, \leq)$ che sia un reticolo è possibile associare una struttura algebrica $(S, \wedge, \vee)$ in cui le due operazioni $\wedge$ e $\vee$ sono definite ponendo

$$x\wedge y = inf\{x,y\} \text{ e } x\vee y = sup\{x,y\}.$$

Tale struttura risulta essere un reticolo, inoltre

$x\leq y \Leftrightarrow x\wedge y = x$.

Proposizione 4.7. Ad ogni struttura algebrica $(S, \vee, \wedge)$ che sia un reticolo possiamo associare una struttura relazionale $(S, \leq)$ definita ponendo $x\leq y$ se e solo se $x\wedge y = x$. Tale struttura risulta essere un reticolo in senso relazionale.

In altre parole è indifferente introdurre i reticoli come strutture relazionali o come strutture algebriche (la questione sarà chiarita nel secondo volume tramite la nozione di categoria). A volte conviene riferirsi alla nozione algebrica a volte a quella relazionale come ad esempio avviene nella seguente definizione.

Definizione 4.8. Un reticolo è detto *distributivo* se

$x\wedge(y\vee z) = (x\wedge y)\vee(x\wedge z)$.

E' detto *limitato* se ammette massimo 1 e minimo 0.

E' subito visto che in un reticolo limitato valgono le seguenti eguaglianze.

$x \vee 0 = 0$; $x \wedge 0 = 0$; $x \wedge 1 = x$; $x \vee 1 = 1$.

La classe più importante di reticoli è forse la seguente.

Definizione 4.9. Un'*algebra di Boole* è una struttura algebrica $(B, \vee, \wedge, -, 0,1)$ dove $(B, \vee, \wedge, 0, 1)$ è un reticolo distributivo limitato ed inoltre valgono le seguenti due uguaglianze

$x \wedge -x = 0$; $x \vee -x = 1$.

Un esempio tipico di algebra di Boole si ottiene considerando un insieme non vuoto S e la struttura $(P(S), \cup, \cap, -, \emptyset, S)$.

5. Gruppi, anelli e campi

Ricordiamo brevemente alcune nozioni elementari di carattere algebrico. La più importante è forse quella di gruppo.

Definizione 5.1. Diciamo che una struttura algebrica $(D, \cdot, ^{-1}, 1)$ è un *gruppo* se:

i) $(x.y)\cdot z = x\cdot(y\cdot z)$ (proprietà associativa)

ii) $x\cdot 1 = x$; $1\cdot x = x$ (1 è elemento neutro)

iii) $x\cdot x^{-1} = 1$; $x^{-1}\cdot x = 1$ (invertibilità).

Un gruppo è detto *commutativo* o *abeliano* se

iv) $x\cdot y = y\cdot x$.

Definizione 5.2. Chiamiamo *anello unitario commutativo* ogni struttura algebrica $(D, +, \cdot, 0, 1)$ tale che:

1) $(D, +, 0)$ sia un gruppo commutativo

2) $(D, \cdot, 1)$ sia una struttura associativa e commutativa con 1 come elemento neutro

3) valga la proprietà distributiva del prodotto rispetto la somma, cioè, per ogni x ed y in D, $(x+y)\cdot z = x\cdot z+y\cdot z$.

Esempio di anello commutativo è quello degli interi relativi. Un altro esempio è quello degli interi modulo un fissato intero m.

Proposizione 5.3. In ogni anello risulta che:

i) $x\cdot 0 = 0$.

ii) $x\cdot(-y) = -x\cdot y$.

iii) $x\cdot(-1)$ è l'opposto di x.

iv) $(-1)^2 = 1$.

Dim. Per provare *i*) osserviamo che per la proprietà distributiva $x\cdot 0 = x\cdot(0+0) = x\cdot 0+x\cdot 0$ da cui, sottraendo da entrambi i membri $x\cdot 0$, si ricava che $0\cdot x = 0$. Per provare *ii*) osserviamo che

$$x\cdot(-y)+x\cdot y = x\cdot(-y+\cdot y) = x\cdot 0 = 0.$$

Le rimanenti proprietà sono ovvie. □

Definizione 5.4. Un anello unitario commutativo $(D, +, \cdot, 0, 1)$ è chiamato *campo* se $(D-\{0\}, \cdot, 1)$ è un gruppo.

Nella teoria degli anelli è importante la nozione di divisore dello zero.

Definizione 5.5. Prende il nome di *divisore dello zero* un elemento $x \neq 0$ tale che esiste $y \neq 0$ tale che $x\cdot y = 0$.

Proposizione 5.6. I divisori dello zero non sono invertibili. Pertanto in un campo non esistono divisori dello zero.

Dim. Sia x un divisore dello zero e supponiamo che esista l'inverso x^{-1} di x. Allora se $y \neq 0$ è tale che $x\cdot y = 0$ si avrebbe che $y = x^{-1}\cdot(x\cdot y) = x^{-1}\cdot 0 = 0$. □

6. Anelli ordinati

Spesso in una struttura algebrica viene definita una relazione d'ordine. In tale caso si richiede che tale relazione "si comporti bene" rispetto alle operazioni. Ad esempio per gli anelli abbiamo la seguente definizione.

Definizione 6.1. Un anello unitario commutativo (D, +, ·, 0, 1) è detto *ordinato* se è definita in D una relazione d'ordine totale $\leq$ *compatibile con le operazioni* cioè tale che

i) $a \leq b \Rightarrow c+a \leq c+b$,

ii) $a \leq b \Rightarrow c \cdot a \leq c \cdot b$ per ogni $c \geq 0$.

Chiamiamo *positivi* gli elementi strettamente maggiori di 0 e *negativi* gli elementi strettamente minori di 0.

Possiamo anche esprimere le condizioni *i*) e *ii*) dicendo che la *traslazione* $f(x) = x+c$ e la *dilatazione* $g(x) = c \cdot x$ con $c > 0$ sono funzioni crescenti. La struttura algebrica degli interi relativi è un esempio di anello ordinato. Poiché ogni campo è un anello, ha senso parlare di campo ordinato. La struttura algebrica definita dai razionali è un tipico esempio di campo ordinato.

Proposizione 6.2. In ogni anello ordinato risulta che:

i) $a \leq b \Rightarrow a\text{-}b \leq 0$; *ii*) $a \leq b \Rightarrow 0 \leq b\text{-}a$.

iii) $b \geq 0 \Leftrightarrow -b \leq 0$; *iv*) $c \geq 0, b \geq 0 \Rightarrow b \cdot c \geq 0$

v) $c \geq 0, a \leq 0 \Rightarrow ac \leq 0$; *vi*) $c \leq 0, b \leq 0 \Rightarrow b \cdot c \geq 0$

vii) $1 \geq 0$, $-1 \leq 0$.

Dim. L'implicazione *i*) si ottiene ponendo $c = -b$ in 1). La *ii*) si ottiene ponendo $c = -a$. Se in i) si pone $a = 0$ allora si ottiene $b \geq 0 \Rightarrow -b \leq 0$. Se in *ii*) si pone $b = 0$ si ottiene $a \leq 0 \Rightarrow 0 \leq -a$. In questo modo la *iii*) è dimostrata. Le rimanenti proprietà si dimostrano in modo analogo. □

Proposizione 6.3. Nel campo dei numeri complessi e nel campo degli interi modulo m non può essere definito un ordinamento che sia compatibile.

Dim. Supponiamo che nel campo dei numeri complessi sia definibile un ordinamento compatibile, allora poiché da *iv*) e *vi*) si ricava che $x^2 \geq 0$ dovremmo avere $-1 = i^2 \geq 0$ in contrasto con la seconda diseguaglianza in *vii*).
Supponiamo che nell'anello degli interi modulo m sia definibile un ordinamento compatibile. Allora dalla diseguaglianza $1 \geq 0$ si

ricava aggiungendo ad entrambi i membri della diseguaglianza che $2\geq 1\geq 0$. Aggiungendo ancora 1 si ottiene che $3\geq 0$ e quindi, procedendo in questo modo, che $-1 = m-1\geq 0$ in contrasto con il fatto che $-1\leq 0$. □

Definizione 6.4. Un elemento x di un anello ordinato si chiama *infinito positivo* se risulta che $x \geq p\cdot 1$ qualunque sia l'intero p. Chiamiamo

- *infinito* negativo l'opposto di un infinito positivo.
- *finito* un elemento che non sia infinito
- *infinitesimo positivo* (*negativo*) un elemento x che sia l'inverso, se esiste, di un elemento infinito positivo (negativo).

L'idea è che x è infinito se, partendo da zero, per quanti passi unitari in avanti si facciano non sia mai possibile superare x. Da notare che x è finito se e solo se esistono p e q tali che $q\cdot 1\leq x\leq p\cdot 1$. Un elemento δ è un infinitesimo positivo se $\delta>0$ e risulta che $1/\delta \geq p\cdot 1$ e quindi $\delta \leq 1/(p\cdot 1)$ per ogni naturale p. Possiamo ora definire la nozione di campo archimedeo.

Definizione 6.5. Un campo ordinato $(D, +, \cdot, 0, 1, \leq)$ si dice *archimedeo* se in esso non esistono elementi infiniti positivi e quindi se verifica il seguente assioma:

$\forall x\exists p\in N(p\cdot 1 \geq x)$ (Assioma di Archimede).

I campi non archimedei sono molto affascinanti perché in essi è possibile sviluppare una teoria degli infiniti e degli infinitesimi.

Infine arriviamo alla più importante proprietà dei campi ordinati, la completezza. Nel seguito dati due sottoinsiemi A e B di un insieme ordinato scriveremo $A \leq B$ per indicare che ogni elemento di A è minore di ogni elemento di B. In tale caso si dice anche che A e B sono *separati*. Chiameremo "*elemento separatore*" di A e B un elemento u tale $A \leq \{u\}\leq B$ cioè un elemento u maggiore di tutti gli elementi di A e minore di tutti gli elementi di B.

Definizione 6.6. Un campo ordinato si dice *completo* se ogni coppia di insiemi separati ammette elemento separatore.

In altre parole un campo ordinato si dice *completo* se verifica l'assioma:

Assioma di completezza: $\forall A \forall B(A \leq B \Rightarrow \exists u(A \leq \{u\} \leq B)$.

Torneremo a parlare di campi ordinati nel prossimo volume quando definiremo i numeri reali dal punto di vista assiomatico

7. La nozione generale di struttura del primo ordine

Abbiamo esaminato molti esempi di strutture relazionali e di strutture algebriche. In questo paragrafo diamo una nozione più generale di struttura matematica.

Definizione 7.1. Una *struttura del primo ordine* è un oggetto matematico del tipo $S = (D, h_1,\ldots, h_n, \mathcal{R}_1,\ldots, \mathcal{R}_m, c_1, \ldots, c_k)$ con

- D insieme non vuoto detto *dominio*
- $h_1, \ldots, h_n$ operazioni
- $\mathcal{R}_1, \ldots, \mathcal{R}_m$ relazioni
- $c_1, \ldots, c_k$ elementi di D.

Se non esistono relazioni allora la struttura prenderà il nome di *struttura algebrica*, se non esistono operazioni prenderà il nome di *struttura relazionale.* I gruppi e gli anelli sono esempi di strutture algebriche. Gli insiemi ordinati sono esempi di strutture relazionali. Gli anelli ordinati (come $\mathbb{Z}$) costituiscono un esempio di struttura algebrica in cui sono presenti sia una relazione che operazioni.

Definizione 7.2. Due strutture $S = (D, h_1, ..., h_n, \mathcal{R}_1, ..., \mathcal{R}_m, c_1, ..., c_k)$ ed $S' = (D', h_1',..., h_n', \mathcal{R}_1',...,\mathcal{R}_m',c_1',..., c_k')$ si dicono *dello stesso tipo* se

i) per ogni i, h_i ed h_i' hanno lo stesso numero di variabili
ii) per ogni j, $\mathcal{R}_j$ e $\mathcal{R}_j'$ si applicano allo stesso numero di elementi.

Ad esempio gli anelli unitari ed i campi sono strutture dello stesso tipo poiché sono forniti di un prodotto, di una somma, di costanti 0 ed 1. I campi ordinati sono di tipo diverso dai campi poiché hanno anche una relazione d'ordine. Nel seguito se $\mathcal{R}$ è una relazione su un insieme D ed $X \subseteq D$, allora la restrizione di $\mathcal{R}$ ad X è la relazione

$$\mathcal{R} \cap X^n = \{(x_1,...,x_n) : x_1 \in X,...,x_n \in X, (x_1,...,x_n) \in \mathcal{R}\}.$$

Denotiamo tale restrizione con $\mathcal{R}/X$.

Definizione 7.3. Una *sottostruttura* di una struttura $S = (D, h_1,...,h_n, \mathcal{R}_1,..., \mathcal{R}_m, c_1,..., c_k)$ è una struttura $S' = (D', h_1',..., h_n', \mathcal{R}_1',..., \mathcal{R}_m', c_1',..., c_k')$, dello stesso tipo di S, tale che:

i) $D' \subseteq D$
ii) $h_i' = h_i/D'$
iii) $\mathcal{R}_i' = \mathcal{R}_i/D'$
iv) $e_i' = e_i$.

Pertanto una sottostruttura di S si ottiene fissando una parte D' di D che sia stabile rispetto alle operazioni $h_1,...,h_n$ e che contenga le costanti $c_1,...,c_k$.

Definizione 7.4. Siano

$$S = (D, h_1,\ldots, h_n, \mathcal{R}_1,\ldots,\mathcal{R}_m, c_1,\ldots,c_k)$$

e

$$S' = (D',h'_1,\ldots,h'_n,\mathcal{R}_1',\ldots,\mathcal{R}_m', c_1',\ldots,c_k')$$

due strutture, chiamiamo *omomorfismo* di S in S' una funzione $f : S_1 \rightarrow S_2$ tale che:

$$(x_1,\ldots,x_n) \in \mathcal{R}_i \Rightarrow (f(x_1),\ldots,f(x_n)) \in \mathcal{R}_i'$$
$$f(h_i(x_1,\ldots,x_n)) = h_i(f(x_1),\ldots,f(x_n))$$
$$f(c_i) = c_i'.$$

Definizione 7.5. Un *epimorfismo* è un omomorfismo suriettivo. Un *isomorfismo* è un omomorfismo biettivo il cui inverso è ancora un omomorfismo. Una *immersione* se è un isomorfismo tra S ed una sottostruttura di S'.

Da notare che la nozione di isomorfismo per le strutture relazionali è leggermente diversa da quella data per le strutture algebriche in cui non si richiede che l'inverso f^{-1} sia ancora un omomorfismo perché ciò accade automaticamente. Invece per le strutture relazionali tale ipotesi è essenziale poiché esistono omomorfismi invertibili il cui inverso non è un omomorfismo. Ad esempio consideriamo due relazioni $\mathcal{R}_1$ ed $\mathcal{R}_2$ rappresentate dai due grafi

$$\begin{tikzcd} a \arrow[r] & b \arrow[d] \\ c & d \arrow[l] \end{tikzcd} \qquad \begin{tikzcd} 1 \arrow[r] \arrow[d] & 2 \\ 3 & 4 \arrow[l] \end{tikzcd}$$

e sia f la funzione definita da $f(a) = 1, f(b) = 2, f(c) = 3, f(d) = 4$. Allora è immediato che f è un omomorfismo invertibile il cui inverso non è un omomorfismo in quanto $1\mathcal{R}_1 3$ ma non è vero che $f^{-1}(1)\mathcal{R}_2 f^{-1}(3)$.

Si pone ora il problema se, data una relazione di equivalenza $\equiv$ in un insieme S in cui siano state definite relazioni ed operazioni se tali relazioni ed operazioni possono essere definite anche nel quoziente $S/\equiv$. Ad esempio, supponiamo che in S sia definita una operazione binaria $+$, allora ha senso, date due classi $X \in S/\equiv$ e $Y \in S/\equiv$, proporre il seguente algoritmo

- prendi un elemento $x \in X$
- prendi un elemento $y \in Y$
- calcola $x+y$
- considera la classe $[x+y]$.

Ma perché una tale definizione funzioni il risultato di un tale algoritmo non deve dipendere dal modo come x ed y sono scelti in X ed Y. In altre parole deve accadere che:

$$x \equiv x' \text{ e } y \equiv y' \Rightarrow x+y \equiv x'+y'.$$

Si perviene allora alla seguente definizione:

Definizione 7.5. Dato un insieme S ed una relazione di equivalenza $\equiv$, diciamo che una operazione n-aria $h : S^n \rightarrow S$ è *compatibile* con $\equiv$ se risulta:

$$x_1 \equiv x'_1, \ldots, x_n \equiv x'_n \Rightarrow h(x_1, \ldots, x_n) \equiv h(x'_1, \ldots, x'_n)$$

Data una relazione n-aria $\mathcal{R}$, diciamo che $\equiv$ è *compatibile con* $\mathcal{R}$ se, per ogni x, x', y, y'

$$(x_1 \equiv x'_1, \ldots, x_n \equiv x'_n) \Rightarrow (x_1, \ldots, x_n) \in \mathcal{R} \Leftrightarrow (x'_1, \ldots, x'_n) \in \mathcal{R}.$$

Data una struttura matematica $S = (D, h_1, \ldots, h_n, \mathcal{R}_1, \ldots, \mathcal{R}_m, c_1, \ldots, c_k)$ chiamiamo *congruenza* una relazione di equivalenza $\equiv$ in D che sia compatibile sia con le operazioni che con le relazioni di tale struttura.

Definizione 7.6. Data una struttura matematica $S = (D,h_1,\ldots,h_n, \mathcal{R}_1,\ldots, \mathcal{R}_m, c_1,\ldots,c_k)$ ed una congruenza $\equiv$, chiamiamo *quoziente di S modulo* $\equiv$, la struttura

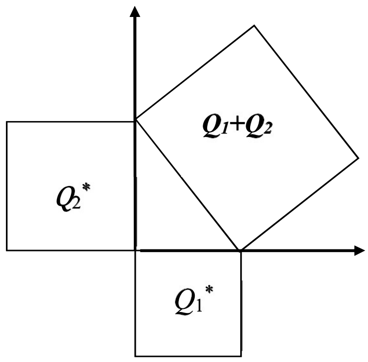
$$S/\equiv\, = (D/\equiv, h_1',\ldots,h_n', \mathcal{R}_1',\ldots, \mathcal{R}_m', [c_1],\ldots,[c_k])$$

dove si è posto

$$h_i'([x_1],\ldots,[x_n]) = [h_i(x_1,\ldots,x_n)],$$
$$\mathcal{R}_i' = \{([x_1],\ldots,[x_n]) : (x_1,\ldots,x_n)\in \mathcal{R}_i\}.$$

Problema: Consideriamo la relazione $\equiv$ in Z definita dal considerare equivalenti due numeri che abbiano gli stessi divisori primi. Ad esempio avremo che $6 \equiv 18$ ma 6 non è equivalente a 730 e la classe [6] contenente 6 è costituita da tutti i numeri che si possono costruire moltiplicando opportunamente 2 e 3, [6]={6, 12, 18,...}. Ancora, il numero 7 è equivalente solo ad una sua potenza e [7] = {7, 49, ...}.

- la relazione $\equiv$ è di equivalenza ?
- la relazione $\equiv$ è compatibile con la moltiplicazione ?
- la relazione $\equiv$ è compatibile con l'addizione ?



Esercizio. Dato l'insieme dei quadrati del piano euclideo, si consideri la seguente operazione Fissato un sistema di assi ortogonali ad ogni coppia di quadrati Q_1 e Q_2 associamo il quadrato Q_1+Q_2 costruito come in figura dove Q_1^* e Q_2^* si ottengono trasportando opportunamente Q_1 e Q_2 sull'asse delle ascisse e su quello delle ordinate. Si ottiene una struttura non commutativa. Dire se l'isometria tra figure è una congruenza ed eventualmente studiare il quoziente relativo. Dire se la corrispondenza che associa ad ogni quadrato la sua area è un omomorfismo.

Esercizio. Consideriamo nell'insieme {1,2,3,4,5,6,7,8,9,10} dei primi dieci numeri naturali, in cui è definita la solita relazione d'ordine le seguenti partizioni

P1 = {{1,2,10}, {3,4,5,6,9}, {7,8}}.
P2 = {{1,2,3,4}, {5,6,7},{8,9,10}}

Tali partizioni determinano a loro volta due relazioni di equivalenza. Dire quale delle due è una congruenza e perché.

8. Anello degli interi modulo *m*.
Per dare un esempio del ruolo delle congruenze e dei relativi quozienti introduciamo ora la nozione di anello degli interi modulo *m*.

Definizione 8.1. Dato un intero $m \geq 2$, chiamiamo *congruenza modulo m,* la relazione $\equiv$ in Z definita ponendo

$x \equiv y \Leftrightarrow x\text{-}y$ è un multiplo di m.

Si dimostra che tale relazione è, appunto, una congruenza nell'anello Z e ciò ci consente di definire il relativo quoziente.

Proposizione 8.2. Dato un intero $m \geq 2$, la struttura quoziente $(Z/\equiv, +, \cdot, [0], [1])$ è un anello commutativo unitario i cui elementi sono le classi $[0], [1], \ldots,[m\text{-}1]$.

Dim. Il fatto che $(Z/\equiv, +, \cdot, [0], [1])$ sia un anello commutativo unitario segue dal fatto che gli assiomi che esprimono l'essere un anello commutativo unitario sono tutti equazioni e che le equazioni si conservano per quoziente. Per provare che gli elementi di Z/m sono $[0], [1], \ldots,[m\text{-}1]$ osserviamo che ogni classe $[n]$ si può rappresentare con un intero n positivo. Infatti nel caso n negativo basta aggiungere ad n un opportuno multiplo di m per ottenere un numero positivo ad esso equivalente. Ad esempio negli interi modulo 5 la classe [-12] coincide con la classe [-12+5+5+5] = [3]. La classe -24 coincide con la classe [-24+5+5+5+5+5] = [1]. D'altra parte sappiamo che, dato un numero naturale n, esistono e sono unici q ed r tali che $n = m \times q + r$ e $0 \leq r \leq m\text{-}1$. Questo significa che esiste un unico r tale che $0 \leq r \leq m\text{-}1$ che sia congruo ad n cioè tale che $[n] = [r]$. □

Si noti che $\equiv$ non è compatibile con l'ordinamento in Z e che quindi non è possibile quozientare Z se lo si intende come struttura ordinata. Infatti sia $\equiv$ ad esempio la congruenza modulo 5. In tali ipotesi $6 \equiv 1$, $5 \equiv 5$ ma pur essendo $1<5$ non è vero che $6<5$.

Proposizione 8.3. L'anello ($\mathbb{Z}/\equiv$, +, ·, [0], [1]) è un campo se e solo se m è primo.

Dim. Se ($\mathbb{Z}/\equiv$, +, ·, [0], [1]) è un campo allora m deve essere primo. Infatti altrimenti, detti x ed y tali che $x \cdot y = m$ risulterebbe $[x] \cdot [y] = [x \cdot y] = [m] = [m]$ con $x<m$ ed $y < m$. Poiché è evidente che $[x] \neq [0]$ e $[y] \neq [0]$, $[x]$ ed $[y]$ risultano essere due divisori dello zero. Ciò è in contrasto con il fatto che un campo è privo di divisori dello zero.

Viceversa, supponiamo che m sia primo e che $[a] \neq [0]$. Consideriamo allora l'applicazione $f([x]) = [a][x]$. Tale applicazione è iniettiva poiché $[a][x] = [a][x']$ comporta che $[a \cdot (x-x')] = [0]$ e quindi che $a \cdot (x-x')$ è un multiplo di m. Essendo per ipotesi m primo e non essendo a un multiplo di m, ciò comporta che sia $x-x'$ un multiplo di m.

Concludiamo che dall'essere f una applicazione iniettiva di un insieme finito in se stesso, f deve essere necessariamente suriettiva. In particolare deve esistere b tale che $f([b]) = [a][b] = 1$ e ciò prova che $[a]$ è invertibile. Poiché ogni elemento non nullo è invertibile, possiamo concludere che $\mathbb{Z}/m$ è un campo. □

Indice Analitico

BIBLIOGRAFIA.

Per la storia della matematica
- B. D'Amore, M. Matteuzzi, *Gli interessi matematici*, Marsilio.
- Benci Vieri, P. Freguglia, *La matematica e l'infinito*, Carocci editore. 2019.
- Biacino Loredana*: Le funzioni elementari: un approccio storico*, Ed. CompoMat, 2009.
- Bottazzini-Freguglia-Rigatelli, *Fonti per la storia della matematica*, Sansoni, 1992.
- G. Lolli, *Da Euclide a Goedel*, Il Mulino, 2004.
- Morris Kline, *La matematica nella cultura occidentale*, Feltrinelli.
- Morris Kline, *Storia del pensiero matematico*, vol I e II, Einaudi, 1972.
- Radice L. L., *L'infinito*, Editori Riuniti.
- Silvio Maracchia, *Storia dell'algebra*, Liguori editore, 2005.
- Attilio Frajese, Attraverso la storia della matematica, Le Monnier 1971.
- Paolo Freguglia, La geometria fra tradizione e innovazione, Bollati-Boringhierei, 1999.

Per quanto riguarda i fondamenti della geometria.
-D. Hilbert, *Fondamenti della geometria*, Feltrinelli, 1970.
- P. Odifreddi, *Divertimento geometrico*: Le origini geometriche della logica da Euclide a Hilbert, Boringhieri, 2003.
-E. Agazzi, D. Palladino, *Le geometrie non euclidee*, Mondadori.
- R. Trudeau, *La rivoluzione euclidea*, Boringhieri, 1991.

Per chi ha interessi verso la filosofia della matematica.
- C. Cellucci, *La filosofia della matematica del Novecento*, Laterza, 2007.
- E. Casari, *La filosofia della matematica del '900*, Sansoni.
- E. Casari, *Questioni di filosofia della matematica*, Feltrinelli.
- G. Lolli, *Filosofia della matematica*: *L'eredità del Novecento*
Lolli G. (2014). *Se viceversa; Trenta pezzi facili e meno facili di Matematica*, Bollati Boringhieri.
- L. Geymonat, *Storia del pensiero filosofico e scientifico*, Garzanti.
- Rudy Rucker, *La mente e l'infinito*, Muzzio, 1991.
- Wang Hao, *Dalla matematica alla filosofia*, Boringhieri, 1984.
-D. R. Hofstadter, Goedel, Escher, *Bach: un' eterna Ghirlanda Brillante*. Il Mulino, 2002.

Per gli aspetti fondazionali della teoria della probabilità

- Costantini D., I fondamenti storico-filosofici delle discipline statistico-probabilistiche, Boringhieri 2004.

Per chi ha interesse al rapporto tra matematica e letteratura

- C. Toffalori, L'aritmetica di Cupido, Guanda Editore 2011.

Per una introduzione dei sistemi numerici.

- M. R. Enea, D. Saeli, *Sistemi Numerici*, ARACNE editrice 2009.
- A. Zanardo, Teorie Assiomatiche, 2010, scaricabile da http://www.math.unipd.it/~azanardo/Metodo/Met_Ass.pdf