

Il paradosso di Banach-Tarski

Alessandro BERARDUCCI*

24 maggio 2018

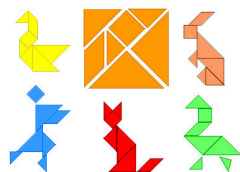
Sommario

L'esposizione del paradosso di Banach-Tarski fornirà l'occasione per rivisitare le nozioni di area e volume sottolineandone alcuni degli aspetti problematici o poco noti e avendo in mente un pubblico di studenti universitari o semplici curiosi.

MSC (2010): 00A09.

1 Congruenze per tagli

Due figure nel piano sono *congruenti per tagli* se possono essere composte a partire da uno stesso insieme di “pezzi” di forma poligonale, ovvero è possibile suddividere l'una in un numero finito di parti di forma poligonale che, opportunamente ridisposte senza alterarne la forma e le dimensioni, formino l'altra figura. Si intende naturalmente che le varie parti non presentino sovrapposizioni al di fuori dei bordi

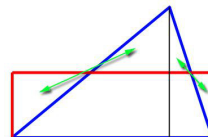


Chi conosca il rompicapo cinese del “Tangram” non avrà difficoltà a riconoscere come le varie figure che possono essere formate con i suoi sette pezzi siano tra loro congruenti per tagli.

Nell'insegnamento dei primi rudimenti di geometria nelle scuole, le congruenze per tagli vengono utilizzate per il calcolo delle aree.

*L'autore ringrazia il supporto del MIUR attraverso il PRIN “Logica, Modelli e Insiemi” e dell'Università di Pisa tramite il PRA “Connessioni fra dinamica olomorfa, teoria ergodica e logica matematica nei sistemi dinamici”.

Possiamo ad esempio ricavare la formula per l'area del triangolo (base per altezza diviso due) sfruttando il fatto che qualsiasi triangolo è congruente per tagli ad un rettangolo con la stessa base ed altezza pari alla metà dell'altezza del triangolo.



Analogamente si dimostra che ogni parallelogramma è congruente per tagli ad un rettangolo. Un attimo in più di riflessione richiede la dimostrazione che un rettangolo è congruente per tagli ad un quadrato, non essendo in questo caso possibile fornire una stima di quanti pezzi siano necessari senza conoscere le proporzioni tra i lati del rettangolo. Componendo queste costruzioni, e triangolando i poligoni, si verifica che ogni poligono è congruente per tagli ad un rettangolo, o anche ad un quadrato. Poiché le congruenze per tagli conservano le aree, ciò fornisce in linea di principio un metodo per determinare l'area di qualsiasi poligono, tuttavia una dimostrazione del fatto che due poligoni sono congruenti per tagli se e solo se hanno la stessa area deve aspettare i lavori di Wallace, Bolyai e Gerwien all'inizio del XIX secolo.

2 Tagli e sottrazioni

Pur non avendo una concezione numerica dell'area, i greci antichi potevano parlare di uguaglianza tra aree e stabilire delle proporzioni. Sapevano ovviamente che figure congruenti per tagli hanno la stessa area, e senza dubbio ammettevano il principio secondo cui *sottraendo aree uguali da aree uguali, si ottengono aree uguali*. Questo ci porta alla seguente definizione, che estende in modo naturale la definizione di congruenza per tagli. Se due figure X ed Y sono congruenti per tagli, e da queste sottraiamo rispettivamente due parti X' di X e Y' di Y che sono anch'esse congruenti per tagli, diremo che le figure risultanti $X - X'$ ed $Y - Y'$ sono *congruenti per tagli e sottrazioni*; si deve in tal caso concludere che hanno la stessa area.

Per illustrare la definizione consideriamo l'enunciato del teorema di Pitagora: dato un triangolo rettangolo, l'area del quadrato costruito sull'ipotenusa è pari alla somma delle aree dei quadrati costruiti sui cateti. Una delle più semplici dimostrazioni, che però non è quella negli Elementi di Euclide, è illustrata in Fig. 1.

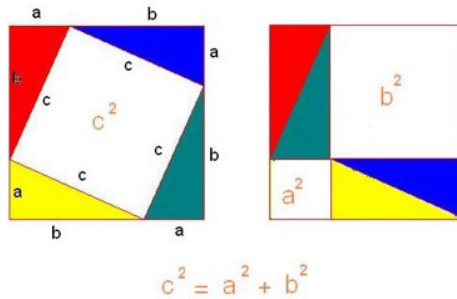


Figura 1: Dimostrazione del teorema di Pitagora per tagli e sottrazioni.

Come si vede dalla figura, sottraendo le aree colorate da ciascuno dei due quadrati grandi di lato $a + b$, si ottiene una dimostrazione del teorema di Pitagora per tagli e sottrazioni.

3 L'assioma di Archimede

Nei Grundlagen der Geometrie, David Hilbert mostra che se due figure sono congruenti per tagli e sottrazioni, lo sono anche per soli tagli. Avendo visto in Fig. 1 una dimostrazione del teorema di Pitagora per tagli e sottrazioni, la Fig. 2 ne illustra una, tra le tante, per soli tagli.

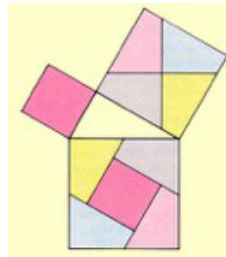


Figura 2: Dimostrazione del teorema di Pitagora per soli tagli.

Dal punto di vista logico-deduttivo può essere interessante osservare che il risultato di Hilbert riposa sull'assioma di Archimede, secondo cui date due grandezze geometriche, in particolare due segmenti, un multiplo intero della più piccola uguaglia o supera la più grande. Per mostrare che esistono geometrie in cui tale assioma non vale ricordiamo che, a partire da Cartesio, la geometria euclidea è riconducibile all'algebra dei numeri reali. Nulla vieta però di costruire in modo del tutto analogo delle geometrie basate su

campi ordinati in cui non vale l'assioma di Archimede¹. In tali geometrie il risultato di Hilbert non è più valido, ma è ancora possibile sviluppare una buona teoria delle aree dei poligoni in modo che due poligoni abbiano la stessa area se e solo se sono congruenti per tagli e sottrazioni². In questo contesto è ancora vero che un rettangolo è congruente per tagli e sottrazioni ad un quadrato, ma potrebbero non bastare i soli tagli. Per giustificare quest'ultima affermazione è sufficiente osservare che i risultati ottenibili senza l'assioma di Archimede producono delle stime uniformi (ricadendo nell'ambito di applicabilità del “teorema di compattezza” per la logica del primo ordine), mentre non è possibile stimare il numero di pezzi necessari a trasformare un rettangolo in un quadrato senza conoscerne il rapporto tra le lunghezze dei lati.

4 Digressione: minimizzare il numero dei pezzi

Ci si può divertire a cercare di minimizzare il numero dei pezzi in una congruenza per tagli. Ad esempio per trasformare un triangolo equilatero in un quadrato bastano quattro pezzi, e per un esagono regolare ne bastano cinque (Fig. 3), ma la suddivisione non è affatto semplice da trovare, o anche solo da verificare dopo averla vista!

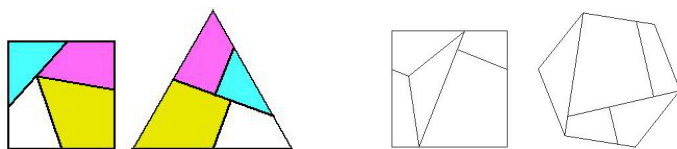


Figura 3: Congruenze per tagli

5 Euclide: l'intero è maggiore della parte

Come semplice esercizio propongo di verificare che le congruenze per tagli godono della *proprietà transitiva*, ovvero se X è congruente per tagli ad Y e Y è congruente per tagli a Z , allora X è congruente per tagli a Z . Se per passare X ad Y bastano n pezzi, e da Y a Z ne bastano m , allora il

¹Non è difficile costruire tali campi, ad esempio possiamo prendere il campo delle funzioni razionali ordinate per crescita asintotica, oppure il campo delle serie di Puiseux, che ha il vantaggio di essere “reale chiuso”.

²Il lettore non avrà difficoltà ad estendere la definizione di congruenza per tagli al caso non archimedeo.

passaggio da X a Z può essere effettuato tramite una suddivisione in $n \cdot m$ pezzi.

Più difficile, soprattutto se non si possiede una buona definizione di area, potrebbe risultare la dimostrazione che una figura X non può essere congruente per tagli ad una sua parte, a meno che la parte coincida con tutto X . La cosa sembra in effetti così ovvia da non richiedere una dimostrazione. Un famoso puzzle dell'enigmista e matematico Henry E. Dudeney (1857-1930) sembra però contraddire questa certezza.

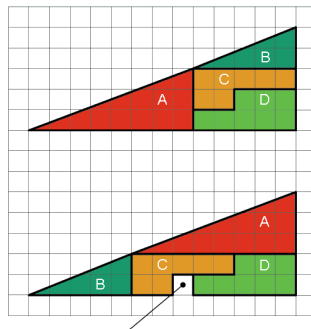


Figura 4: Dove è finito il quadratino mancante?

Il puzzle, illustrato in Fig. 4, può lasciare a prima vista sconcertati in quanto sembra mettere in crisi il concetto stesso di area, insieme alla nozione euclidea secondo cui l'intero è maggiore della parte. C'è però un inganno che invito a scoprire, e in effetti si dimostra che simili "sparizioni" per effetto di un semplice movimento di pezzi non possono capitare. Vedremo a questo proposito che la terza dimensione riserverà delle sorprese.

6 Volume dei solidi: la piramide

L'obiettivo finale sarà quello di capire se per i solidi valgono risultati analoghi a quelli già visti per le figure piane, ma per acquisire un po' di dimestichezza con l'argomento iniziamo con il calcolare alcuni volumi. Esattamente come un quadrato può essere diviso in due triangoli rettangoli, un cubo può essere diviso in tre piramidi identiche a base quadrata, come in Fig. 5.

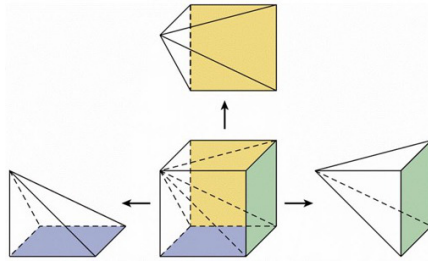


Figura 5: Volume della piramide

Ne deduciamo che il volume di ciascuna delle tre piramidi è dato dalla formula “area di base per altezza diviso tre”, ed è significativo che questi conti possano essere fatti senza possedere una definizione precisa di volume, bensì utilizzando alcune proprietà che evidentemente assumiamo in forma implicita.

7 Principio di Cavalieri

Se la piramide ha una forma diversa possiamo determinarne il volume tramite il cosiddetto “principio di Cavalieri”. Per illustrare il principio, consideriamo due solidi, o nel caso specifico due piramidi, come in Fig. 6. Supponiamo che si possano posizionare i due solidi nello spazio in modo tale che, sezionandoli con piani paralleli ad un piano dato, le rispettive sezioni abbiano sempre la stessa area. Secondo il principio che prende il nome da Bonaventura Cavalieri (1598-1647) i due solidi devono necessariamente avere lo stesso volume.

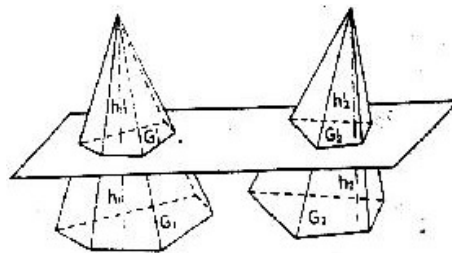


Figura 6: Principio di Bonaventura Cavalieri, 1598-1647

Si dimostra in questo modo che due piramidi con la stessa area di base e la stessa altezza hanno lo stesso volume, e siccome abbiamo già calcolato il

volume di alcune piramidi di forma particolare, possiamo facilmente dedurre che il volume di una generica piramide si calcola di nuovo con la medesima formula: “area di base per altezza diviso tre”.

Il principio di Cavalieri può essere dimostrato facendo ricorso al metodo di esaustione di Archimede nel modo seguente. Consideriamo il caso specifico delle due piramidi nella figura e immaginiamo di sostituire le sezioni bidimensionali con delle sottili “fettine” solide, ciascuna ritagliata da due piani paralleli molto vicini tra loro. Il volume di una di queste fettine può essere approssimato dall’area della sua base per la rispettiva altezza. Dico “approssimato”, anziché calcolato esattamente, perché la formula “area di base per altezza” presuppone che le fettine siano prismi anziché piramidi tronche. Sommando i contributi dati dai vari prismi, si ottiene il volume di due strutture a gradini, tipo “ziggurat”, che approssimano le piramidi. Per come sono state costruite, le due strutture a gradini hanno lo stesso volume, e facendo fettine sempre più sottili il loro volume approssima quello delle rispettive piramidi, con un errore che può essere reso piccolo quanto si vuole. Se ne deve necessariamente concludere che le due piramidi hanno lo stesso volume. Su analoghi processi di approssimazione si basa la moderna teoria dei volumi e degli integrali.

8 Il terzo problema di Hilbert

Estendiamo dal piano allo spazio il concetto di congruenza per tagli, semplicemente richiedendo che i pezzi della scomposizione siano poliedrali anziché poligonali (e sempre richiedendo che non vi siano sovrapposizioni delle parti interne). Possiamo allora chiederci se due poliedri con lo stesso volume siano sempre congruenti per tagli. In contrasto con quanto avveniva nel caso poligoni, la risposta è però negativa. Nel 1902, risolvendo il terzo dei famosi ventitre problemi posti da Hilbert al congresso internazionale dei matematici di Parigi nel 1900, Max Dehn ha infatti dimostrato la cosa seguente.

Un cubo non è congruente per tagli ad un tetraedro, ovvero non è possibile suddividere un cubo in un numero finito di poliedri che possano essere ridisposti in modo da formare un tetraedro.

La dimostrazione si basa sul fatto che è possibile associare ad ogni poliedro un invariante algebrico (una specie di numero, ma un po’ più complicato degli usuali numeri che tutti conoscono) in modo tale che due polietri congruenti per tagli abbiano lo stesso invariante. A questo punto basta far vedere che per un cubo l’invariante di Dehn è uguale a zero mentre per un tetraedro è diverso da zero, ed il gioco è fatto.

Per chi voglia approfondire, l'invariante di Dehn è un elemento del prodotto tensore $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}/\pi\mathbb{Z}$, ed è definito come la somma $\sum_i \ell_i \otimes \theta_i$, estesa a tutti i lati del poliedro, essendo ℓ_i la lunghezza del lato i e θ_i il corrispondente angolo diedrale. Una dimostrazione diversa si può trovare in [1].

9 Il principio di continuità

Alla luce del risultato di Dehn, potrebbe sorgere il dubbio che esista un cubo dello stesso volume di un tetraedro. Una risposta positiva si ottiene in base al “principio di continuità”: se teniamo fisso il tetraedro e ingrandiamo progressivamente un cubo inizialmente molto piccolo, dobbiamo necessariamente passare da cubi di volume decisamente inferiore a cubi di volume decisamente superiore a quello del tetraedro, e per continuità il volume dovrà passare per tutte le misure intermedie, assumendo anche esattamente quella del tetraedro.

10 Equiscomposizioni

È naturale porsi la domanda se il teorema di Dehn continui a valere rilassando la richiesta che i pezzi della congruenza per tagli siano poliedri. Diamo quindi la seguente definizione.

Due spazi metrici X, Y sono equiscomponibili, se ammettono partizioni finite $X = X_1 \cup \dots \cup X_n$ e $Y = Y_1 \cup \dots \cup Y_n$ tali che X_i ed Y_i si corrispondano tramite un'isometria per ogni $i = 1, \dots, n$.

Nel caso X ed Y siano dei solidi nello spazio, possiamo modificare la definizione limitandoci alle isometrie che preservano l'orientazione, cosa che è possibile fare ai fini del teorema di Banach-Tarski. In questo caso la richiesta nella definizione equivale a dire che è possibile spezzare X ed Y in un numero finito di parti in modo che le parti corrispondenti X_i ed Y_i siano identiche nella forma e nelle dimensioni, salvo essere posizionate in modo diverso nello spazio, di modo che un movimento rigido possa portare X_i e Y_i a sovrapporsi esattamente.

Le equiscomposizioni sono molto simili alle congruenze per tagli, ma con due importanti differenze: da un lato nelle equiscomposizioni si ammettono pezzi di forma arbitraria (non necessariamente poliedrali); dall'altro però si insiste sul fatto che i pezzi formino una partizione, ovvero siano del tutto disgiunti tra loro, mentre nelle congruenze per tagli si richiedeva solamente che fossero le loro parti interne ad essere disgiunte, ammettendo sovrapposizioni dei bordi.

Si può dimostrare che se X ed Y sono congruenti per tagli sono anche equiscomponibili, ovvero si possono eliminare le sovrapposizioni lungo i bordi, ma al prezzo di usare pezzi più complicati. Come la congruenza per tagli, anche la equiscomponibilità è una relazione di equivalenza.

11 Teorema di Banach-Tarski

Abbiamo visto che un cubo e un tetraedro dello stesso volume non sono congruenti per tagli, ed è naturale chiedersi se siano equiscomponibili. La buona notizia è che in effetti lo sono, il che farebbe pensare che le equiscomposizioni siano uno strumento migliore rispetto alle congruenze per tagli. Quella cattiva, abbastanza stupefacente, è che lo sono anche se non hanno lo stesso volume! Ciò dovrebbe ricordare il puzzle di Dudeney precedentemente visto in Fig. 4, salvo che ora non c'è inganno. Il teorema di Banach-Tarski afferma, in generale, che *due solidi qualsiasi con parte interna non vuota sono equiscomponibili*.

Una palla del diametro di una biglia è dunque equiscomponibile ad una del diametro del sole: possiamo suddividere l'una in un numero finito di parti disgiunte che, ridisposte in altro modo nello spazio senza alterarne forma e dimensioni, formino l'altra. In definitiva, mentre con le congruenze per tagli non ce la si fa anche quando ce la si dovrebbe fare (teorema di Dehn), con le equiscomposizioni, ce la si fa anche quando non ce la si dovrebbe fare (teorema di Banach-Tarski). I due risultati mostrano che, volendo rendere precisa l'idea intuitiva di "avere lo stesso volume", non possiamo far ricorso né all'esistenza di una congruenza per tagli, né all'esistenza di un'equiscomposizione.

A differenza di altri paradossi, come quello di Zenone o quello del mentitore, destinati a rimanere interrogativi problematici, quello di Banach-Tarski non è in realtà un paradosso, bensì un teorema matematico, nonostante sia così incredibile da apparire paradossale.

12 Cos'è il volume?

A questo punto è inevitabile chiedersi quale sia la corretta definizione matematica di volume. Questo è un argomento importante che però non posso approfondire, limitandomi ad accennare al fatto che la definizione si basa sul concetto di approssimazione (limiti, integrali), come nella giustificazione che abbiamo dato del principio di Cavalieri tramite il metodo di esaustione di Archimede. Una delle conseguenze del teorema di Banach-Tarski (ma

segue anche da altri risultati precedenti), è che questo processo di approssimazione è applicabile solo ad una certa classe di insiemi, i cosiddetti insiemi Lebesgue-misurabili. Dobbiamo rassegnarci al fatto che a certi insiemi particolarmente complicati non sarà possibile assegnare un volume.

13 Strategia della dimostrazione

La seconda parte di questo intervento richiede una certa concentrazione, ma invito il lettore a non scoraggiarsi; anche una lettura superficiale può dare i suoi frutti. La dimostrazione (a grandi linee) del teorema di Banach-Tarski sarà infatti l'occasione, o il pretesto, per introdurre in modo informale qualche concetto importante della matematica. Le varie sezioni sono monotematiche e solo alla fine sarà chiaro, almeno spero, come amalgamare i vari ingredienti.

Per dimostrare il teorema di Banach-Tarski ci si riconduce a mostrare che un solido sferico è “duplicabile”, ovvero è equiscomponibile con due copie identiche di sé stesso (si può dimostrare che bastano cinque pezzi!). Per raggiungere questo obiettivo faremo prima vedere che la superficie sferica privata di un certo sottoinsieme numerabile è duplicabile (Teorema di Hausdorff).

I pezzi coinvolti nelle duplicazioni sono così complicati che non è possibile disegnarli o farsene una semplice immagine visiva. Tuttavia la loro esistenza concettuale riposa su principi che oggi (dopo varie esitazioni iniziali) sono comunemente accettati dai matematici, tra cui il cosiddetto “assioma della scelta”; esso servirà per scegliere un insieme di rappresentanti delle orbite di certe azioni gruppal. Se si trascurano i bordi, vi sono anche versioni del paradosso che non usano l'assioma della scelta (teorema di Dougherty e Foreman).

14 L'intero è veramente maggiore della parte?

Uno dei più noti paradossi dell'infinito può essere illustrato dal cosiddetto “albergo di Hilbert”. Esso ha infinite stanze, tante quante i numeri naturali $0, 1, 2, 3, \dots$, e sono tutte occupate. All'arrivo di un nuovo cliente l'albergatore riesce ad alloggiarlo con dei semplici spostamenti di stanza. Come fa? La risposta non è difficile: visto che non vi è un'ultima stanza, basta chiedere a ciascun cliente di spostarsi nella stanza successiva (dalla 0 alla 1, dalla 1 alla 2, dalla 2 alla 3, eccetera), in modo che la stanza 0 si liberi e possa accogliere il nuovo cliente.

15 Un primo paradosso geometrico

Con la stessa idea dell'albergo di Hilbert possiamo dimostrare che un cerchio (la parte di piano delimitata da una circonferenza) è equiscomponibile al cerchio stesso privato di uno dei raggi. Più precisamente un cerchio è scomponibile in due pezzi P e Q che possono essere ruotati in modo da formare il cerchio stesso privato di un raggio.

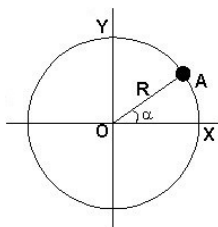


Figura 7: Un cerchio e il cerchio meno un raggio sono equiscomponibili.

Si usa la tecnica dell'albergo di Hilbert. Scegliamo l'angolo α in modo che $\alpha, 2\alpha, 3\alpha, 4\alpha, \dots$ siano tutti distinti (basta scegliere α in modo che α/π sia irrazionale). Sia Q l'insieme formato dai raggi che hanno coordinata angolare pari ad uno degli $n\alpha$ (n intero positivo), e sia P la parte restante del disco (conveniamo che il centro del cerchio sia in P). Ora risistemiamo i pezzi P e Q come segue: P lo lasciamo fermo, mentre Q lo ruotiamo di un angolo α . In tal modo il raggio $n\alpha$ si sposta in posizione $(n+1)\alpha$ e il raggio di angolo α rimane scoperto.

Questo paradosso è meno sorprendente di quello di Banach-Tarski perché il cerchio privato di un raggio ha la stessa area di tutto il cerchio, mentre nel paradosso di Banach-Tarski si riescono ad alterare i volumi. Tuttavia abbiamo fatto un primo passo.

16 Il concetto di gruppo

Uno dei concetti più importanti dell'algebra moderna è quello di gruppo. Per i nostri fini possiamo concentrarci sui gruppi di "movimenti", ad esempio le mosse che è possibile effettuare nel tentativo di risolvere quell'ingegnoso rompicapo noto come "cubo di Rubik".

Altri esempi di gruppi sono dati dalle isometrie del piano e dello spazio. Ricordiamo che le isometrie sono le trasformazioni che non alterano le di-

stanze, e nel caso del piano includono le rotazioni intorno ad un punto, le traslazioni, e i ribaltamenti intorno ad un asse.

Se ci restringiamo ai movimenti che conservano certe figure o simmetrie otteniamo un sottogruppo, ovvero un gruppo più piccolo incluso in quello di partenza. Possiamo ad esempio considerare il sottogruppo consistente di quei movimenti del piano che portano il motivo raffigurato in Fig. 8 (esteso all'infinito in tutte le direzioni) a sovrapporsi a sé stesso.

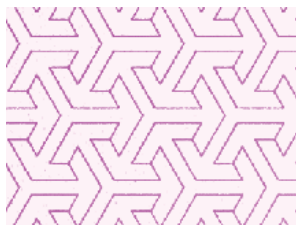


Figura 8: Simmetrie e sottogruppi

Se x ed y sono elementi di un gruppo, anche la loro composizione, che indicheremo con la notazione moltiplicativa xy , è un elemento del gruppo, e corrisponde a fare prima³ il movimento x e poi y . Il movimento nullo, che lascia tutto fermo, viene chiamato *elemento neutro* del gruppo ed è generalmente indicato con il simbolo 1 , oppure con la lettera e . Si assume che ogni elemento x di un gruppo possenga un *inverso* x^{-1} , anch'esso appartenente al gruppo, definito come quell'elemento tale che $xx^{-1} = 1$. In altre parole x^{-1} coincide con il movimento x effettuato all'incontrario, di modo che fare prima x e poi x^{-1} riporti alla posizione di partenza.

17 Il grafo di Cayley di un gruppo

Consideriamo ora il gruppo D_3 delle simmetrie del triangolo, consistente nei sei movimenti rigidi che portano un triangolo a sovrapporsi a sé stesso (tre rotazioni intorno al punto centrale e tre ribaltamenti lungo gli assi di simmetria). Esso può essere generato dalla rotazione di 120 gradi nel senso della vite levogira e dal ribaltamento intorno all'asse verticale di simmetria, indicati rispettivamente con le lettere a e b come nelle Figg. 9 e 10, essendo le altre quattro simmetrie ottenibili come composizione di queste due.

³Seguiamo la convenzione che i movimenti si effettuano nell'ordine in cui si scrivono da sinistra a destra. Altri autori seguono la convenzione opposta.

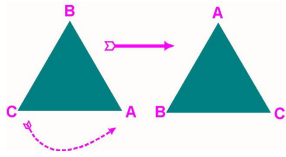


Figura 9: $a =$ rotazione

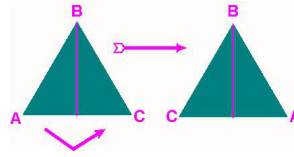


Figura 10: $b =$ ribaltamento

In generale dato un gruppo, e specificato un insieme di generatori, possiamo considerare il cosiddetto *grafo di Cayley* del gruppo, il quale rappresenta in modo visivo quali relazioni sussistano tra i generatori. Ad esempio nel gruppo D_3 valgono le relazioni $a^3 = 1$ (ruotare tre volte riporta alla posizione di partenza), $b^2 = 1$ (ribaltare due volte equivale a non fare nulla) e $ba = a^2b$ (ribaltare e poi ruotare equivale a ruotare due volte e poi ribaltare). Queste relazioni, e quelle che ne derivano, sono leggibili nel grafo di Cayley di D_3 rappresentato in Fig. 11.

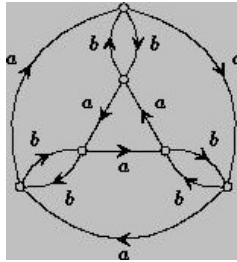


Figura 11: Grafo di Cayley di D_3

I sei nodi del grafo sono in corrispondenza biunivoca con gli elementi del gruppo e rappresentano le sei possibili posizioni ABC , CAB , BCA , CBA , BAC e ACB che può assumere il triangolo (dove A, B, C sono i vertici), mentre le frecce rappresentano l'azione dei generatori e portano da una posizione a quella risultante dopo aver effettuato il movimento. Ad esempio la relazione $a^3 = 1$ corrisponde al fatto che seguire per tre volte la freccia "a" a partire da qualsiasi posizione riporta alla posizione di partenza, mentre la relazione $ba = a^2b$ rispecchia il fatto che, partendo da qualsiasi nodo e seguendo le frecce b e a (in quest'ordine), si arriva nello stesso nodo che seguendo le frecce a, a e b . Associando ad uno dei nodi l'elemento neutro del gruppo, si ottiene una corrispondenza biunivoca tra i nodi del grafo di Cayley e gli elementi del gruppo.

Abbiamo fin qui parlato di gruppi di movimenti, che sono quelli che maggiormente ci interessano, ma è possibile estendere in modo naturale la definizione di grafo di Cayley a gruppi astratti.

18 Gruppi liberi

Due elementi a e b di un gruppo G si dicono *indipendenti* se il sottogruppo $\langle a, b \rangle$ da essi generato è *libero*. Ciò significa che non vi sono relazioni tra i generatori se non quelle obbligatorie che valgono in tutti i gruppi, quale potrebbe ad esempio essere $xx^{-1} = 1$. Il gruppo D_3 precedentemente visto non è libero in quanto vi sono le relazioni $a^3 = 1, b^2 = 1$ e $ba = a^2b$. Analogamente, un gruppo commutativo con almeno due elementi non è libero, in quanto tra due suoi elementi x ed y vi è sempre almeno la relazione di commutatività $xy = yx$, che possiamo scrivere anche nella forma $xyx^{-1}y^{-1} = 1$. In generale le relazioni si esprimono nella forma $w = 1$, dove w è una *parola ridotta*, ovvero una composizione dei generatori e dei loro inversi in cui un generatore non sia mai affiancato al suo inverso (altrimenti potremmo cancellare quella coppia, come in $yx^{-1}x = y$).

La Fig. 12 illustra il grafo di Cayley di un gruppo libero $G = \langle a, b \rangle$ con due generatori a e b .

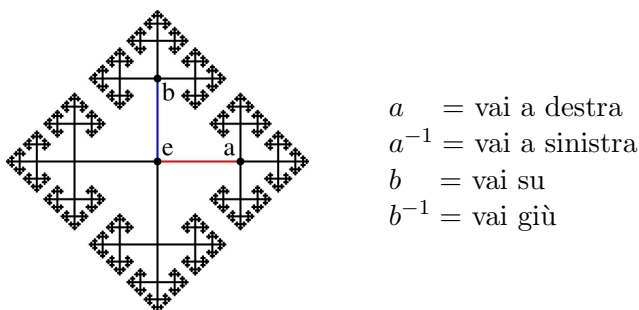
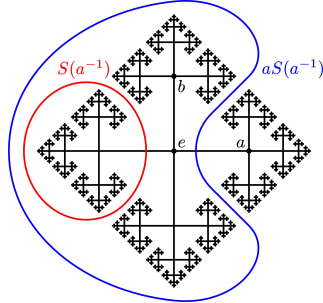


Figura 12: Grafo di Cayley del gruppo libero con due generatori

L'assenza di relazioni si riflette nel fatto che tra due nodi qualsiasi vi sia un solo cammino che li congiunge: se ve ne fossero due, otterremmo la relazione non banale $w = v$ dove w e v sono le parole ridotte corrispondenti ai due cammini. Come già detto, i nodi del grafo di Cayley corrispondono biunivocamente agli elementi del gruppo. Indicando con $S(a), S(b), S(a^{-1})$ e $S(b^{-1})$ i sottoinsiemi del gruppo libero $G = \langle a, b \rangle$ dati dalle parole ridotte che iniziano per a, a^{-1}, b e b^{-1} rispettivamente, si vede che essi corrispondono ai

quattro “petali” della Fig. 12. Ad esempio $abab^{-1}$ appartiene a $S(a)$, mentre la parola vuota corrisponde al centro delle figura e rappresenta l’elemento neutro del gruppo.

In analogia con le equiscomposizioni geometriche, possiamo definire delle equiscomposizioni gruppali: il ruolo dei movimenti rigidi sarà preso dalle traslazioni a sinistra. In Fig. 13 abbiamo indicato con $aS(a^{-1})$ l’insieme degli elementi che si ottengono da $S(a^{-1})$ tramite una traslazione a sinistra per l’elemento a . Ad esempio l’elemento b appartiene ad $aS(a^{-1})$ in quanto possiamo scriverlo nella forma $b = aa^{-1}b$, ovvero moltiplicando a per l’elemento $a^{-1}b$ di $S(a^{-1})$. In generale si vede che ogni elemento del gruppo, o appartiene a $S(a)$ oppure a $aS(a^{-1})$, cosicché $G = S(a) \cup aS(a^{-1})$. Abbiamo così dimostrato che G è equiscomponibile tramite traslazioni a sinistra all’insieme $A = S(a) \cup S(a^{-1})$ dato dall’unione dei petali est-ovest. Nello stesso modo si mostra che G è equiscomponibile per traslazioni a sinistra all’insieme $B = S(b) \cup S(b^{-1})$, dato dall’unione dei petali nord-sud, in quanto possiamo scriverlo nella forma $G = S(b) \cup bS(b^{-1})$. Tutto ciò è illustrato in Fig. 13, dove tutte le unioni sono tra parti disgiunte.



$$\begin{aligned}
 G &= 1 \cup S(a) \cup S(a^{-1}) \cup S(b) \cup S(b^{-1}) \\
 &= S(a) \cup aS(a^{-1}) \\
 &= S(b) \cup bS(b^{-1})
 \end{aligned}$$

Figura 13: Duplicazione del gruppo libero

Le due scomposizioni $G = S(a) \cup aS(a^{-1}) = S(b) \cup bS(b^{-1})$ mostrano che il gruppo libero è “duplicabile”: da ciascuna coppia di petali opposti possiamo ricreare l’intero gruppo per equiscomposizioni gruppali. Il prossimo obiettivo è trasformare queste equiscomposizioni gruppali in equiscomposizioni geometriche.

19 Rotazioni indipendenti nello spazio

Per dimostrare il teorema di Banach-Tarski a noi interesserà trovare sottogruppi liberi all’interno di gruppi di rotazioni. Nel piano ciò non è possibile

in quanto il gruppo di rotazioni del piano è commutativo (fissata un'origine, ruotare un piano di un angolo α intorno all'origine e poi di un angolo β è la stessa cosa che ruotare prima di β e poi di α). In generale non è possibile trovare gruppi liberi con due generatori all'interno di un gruppo "risolubile", come ad esempio il gruppo delle isometrie del piano. Vedremo però che nella terza dimensione le cose cambiano.

Le rotazioni nello spazio coincidono con le trasformazioni che fissano l'origine e conservano le distanze e l'orientazione. Ciascuna rotazione è individuata da un asse orientato passante per l'origine e un angolo di rotazione intorno all'asse (conveniamo che l'angolo indichi l'ampiezza della rotazione nel senso di una vite levogira). Componendo due rotazioni si ottiene un'altra rotazione, in generale intorno ad un terzo asse, che dipenderà anche dall'ordine in cui si effettuano le rotazioni, onde la non commutatività. Una sfera centrata nell'origine viene portata in sé stessa da una rotazione dello spazio, cosicché il gruppo delle rotazioni dello spazio agisce sulla sfera e coincide in effetti con il gruppo delle rotazioni della sfera in sé stessa. Tale gruppo si chiama $SO(3, \mathbb{R})$.

Hausdorff ha dimostrato che in $SO(3, \mathbb{R})$ esistono due elementi indipendenti. Possiamo darne una dimostrazione desunta da un articolo di Terence Tao pubblicato sulla sua pagina web. Consideriamo le rotazioni rappresentate dalle matrici a e b definite come segue

$$a = \frac{1}{5} \begin{pmatrix} 3 & 4 & 0 \\ -4 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \quad b = \frac{1}{5} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & -3 \end{pmatrix}$$

(la presenza del fattore $1/5$ garantisce che la matrice sia ortogonale, con determinante 1, e rappresenti dunque una rotazione). Supponiamo per assurdo che vi sia una relazione non banale che leghi a e b . Come abbiamo visto, una tale relazione può essere espressa nella forma $w = 1$, dove 1 è il movimento nullo, rappresentato dalla matrice identità, e w è una parola ridotta in a, b, a^{-1} e b^{-1} , ovvero una composizione non banale di queste quattro matrici. La corrispondente composizione delle matrici $5a, 5b, 5a^{-1}$ e $5b^{-1}$ (non più ortogonale) è una matrice a coefficienti interi della forma $5^n I$ dove I è la matrice identità e n è il numero di elementi che stiamo componendo. Per raggiungere un assurdo consideriamo l'insieme dei quattro vettori seguenti:

$$\left\{ \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -4 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} \right\}.$$

Ragionando modulo 5, ciascuna delle matrici $5a, 5b, 5a^{-1}$ e $5b^{-1}$ manda ciascuno dei quattro vettori in un altro vettore dello stesso insieme. La stessa cosa deve allora avvenire per $5^n I$, essendo una loro composizione. Questo è assurdo in quanto se applichiamo $5^n I$ ad un qualsiasi vettore otteniamo un vettore i cui coefficienti sono congrui a zero modulo 5.

20 Orbita di un punto

Prima di proseguire con la dimostrazione dobbiamo spiegare il concetto di orbita rispetto ad una azione gruppale. Ci limiteremo al caso che ci interessa, dato dai gruppi di rotazione. Dato un sottogruppo G di $SO(3, \mathbb{R})$ e un punto P sulla superficie sferica, l'*orbita* di P consiste nell'insieme dei punti $GP = \{gP \mid g \in G\}$ in cui può andare a finire P sotto l'azione di una delle rotazioni di G .

Consideriamo ad esempio il sottogruppo G dato da tutte le possibili rotazioni intorno all'asse verticale. In questo caso l'orbita di P è il *parallelo* che passa per P , come in Fig. 14, e si riduce al solo punto P se P è uno dei due poli.

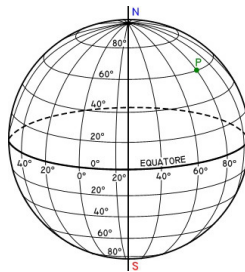


Figura 14: I paralleli sono le orbite rispetto alle rotazioni lungo l'asse di rotazione

Un *insieme di rappresentanti* per le orbite è un insieme che contiene uno ed un solo punto per ogni orbita.

Nel nostro esempio un possibile insieme di rappresentanti è dato da un qualsiasi meridiano. Se scegliamo un sottogruppo di $SO(3, \mathbb{R})$ più complicato, può non essere altrettanto semplice “disegnare” un insieme di rappresentanti per le orbite, ma la sua esistenza concettuale è garantita dall'assioma della scelta e deve la sua giustificazione intuitiva all'idea che possiamo idealmente “scegliere” in modo arbitrario un punto da ciascun orbita.

21 Duplicazione della sfera

Ricordiamo che un insieme è numerabile se può essere messo in corrispondenza biunivoca con i numeri interi. In alcuni dei lavori matematici più importanti del XIX secolo, Georg Cantor ha dimostrato che non tutti gli insiemi infiniti hanno lo stesso numero di elementi, o più precisamente la stessa “cardinalità”. Gli insiemi infiniti più piccoli sono gli insiemi numerabili, ma i punti di un segmento geometrico hanno la cardinalità del continuo e non possono essere messi in corrispondenza biunivoca con i numeri interi.

Il teorema di Hausdorff afferma che la superficie sferica, privata di un opportuno sottoinsieme numerabile, è equiscomponibile con due copie di sé stessa.

Per dimostrare il teorema consideriamo il sottogruppo $G = \langle a, b \rangle$ di $\text{SO}(3, \mathbb{R})$ generato da due rotazioni indipendenti a e b , e osserviamo che tale sottogruppo, i cui elementi sono in corrispondenza con i nodi del grafo di Cayley in Fig. 12, è numerabile. Ciascun elemento g di G è una rotazione della sfera di un certo angolo intorno ad un opportuno asse. Indicando con S^2 la superficie sferica, in corrispondenza a g vi saranno due punti “eccezionali” di S^2 , dati dall’intersezione dell’asse di rotazione con S^2 . Se g è diverso dal movimento nullo, questi saranno gli unici due punti lasciati fermi da g . Ora sia E il sottoinsieme numerabile dato da tutti i punti eccezionali di S^2 al variare di g in G . L’insieme dei punti della superficie sferica ha cardinalità del continuo, quindi la maggior parte dei punti non sono eccezionali.

Se x è un punto non eccezionale di S^2 , la sua orbita per l’azione del sottogruppo G è in corrispondenza biunivoca con G stesso tramite la funzione che manda $g \in G$ nel punto $gx \in S^2$ (la rotazione g applicata al punto x). Dato che il gruppo libero G è “duplicabile” tramite traslazioni a sinistra, ne segue che l’orbita $Gx \subseteq S^2 - E$ di x è duplicabile tramite rotazioni. Più precisamente, esistono due sottoinsiemi disgiunti e non vuoti A e B del gruppo G (corrispondenti ai petali est-ovest e a quelli nord-sud in Fig. 13) tali che l’orbita Gx è equiscomponibile con ciascuno dei suoi due sottoinsiemi disgiunti Ax e Bx (e avanzerà anche un punto, visto che $Gx = Ax \cup Bx \cup \{x\}$). Abbiamo così duplicato l’orbita di x ; poiché $S^2 - E$ è un’unione di orbite, possiamo analogamente duplicare tutto $S^2 - E$ dopo aver scelto un insieme di rappresentanti per le orbite.

Scendendo più nel dettaglio, sia $M \subseteq S^2 - E$ un insieme di rappresentanti per le orbite dei punti non eccezionali. La sfera privata dei punti eccezionali, ovvero $S^2 - E$, coincide con l’unione $GM = \bigcup_{x \in M} Gx$ delle orbite dei punti x di M , ed è equiscomponibile con ciascuno dei suoi due sottoinsiemi disgiunti

AM e BM , ovvero da ciascun di questi due sottoinsiemi possiamo ricreare l'intero insieme $S^2 - E$ tramite equiscomposizioni. Abbiamo finalmente dimostrato il teorema di Hausdorff: $S^2 - E$ è duplicabile.

Per dimostrare che S^2 stesso è duplicabile, basta mostrare che $S^2 - E$ è equiscomponibile all'intero S^2 . Si utilizza a tal fine la tecnica dell'albergo di Hilbert, la stessa che abbiamo utilizzato per mostrare che il cerchio è equiscomponibile con il cerchio meno un raggio (Fig. 7). In quel caso avevamo scelto una rotazione di un'opportuno angolo per "riassorbire" il raggio facendolo sparire. Analogamente possiamo riassorbire l'insieme E dei punti eccezionali tramite un'opportuna rotazione ρ intorno ad un asse passante per il centro della sfera e tale che gli insiemi $E, \rho(E), \rho^2(E), \rho^3(E), \dots$ siano tutti disgiunti. L'esistenza di una tale rotazione è garantita dal fatto che E è numerabile, cioè non molto grande, anche se infinito.

Dalla duplicabilità di S^2 si ottiene infine facilmente la duplicabilità della sfera solida meno il suo centro (basta considerarla come unione di superfici sferiche concentriche). La duplicabilità dell'intera sfera solida si ottiene riassorbendo il centro con una rotazione non periodica per un asse non passante per il centro medesimo.

22 Teorema di Banach-Schröder-Bernstein

Il teorema di Schröder-Bernstein afferma che dati due insiemi X ed Y , se esiste una funzione iniettiva da X ad Y e una iniettiva da Y ad X , allora vi è anche una funzione biunivoca tra X ed Y . Per dimostrarlo si usa la tecnica dell'albergo di Hilbert per far scorrere certi punti lungo certe successioni. Il teorema facilita notevolmente il compito di dimostrare l'esistenza di funzioni biunivoche. Ad esempio una sua immediata conseguenza è che esiste una corrispondenza biunivoca tra un segmento aperto e un segmento chiuso della retta reale, essendo evidentemente molto semplice trovare funzioni iniettive in entrambi i versi.

Un teorema simile è stato dimostrato da Banach per le equiscomposizioni e viene chiamato teorema di Banach-Schröder-Bernstein: *se X è equiscomponibile ad un sottoinsieme di Y ed Y è equiscomponibile ad un sottoinsieme di X , allora X ed Y sono tra loro equiscomponibili.*

Avendo dimostrato che una sfera solida è duplicabile, e facendo uso del risultato appena enunciato, è facile dimostrare che due solidi qualsiasi con parte interna non vuota sono equiscomponibili, completando la dimostrazione del teorema di Banach-Tarski.

23 Ruolo dell'assioma della scelta

Nel teorema di Banach-Tarski, l'assioma della scelta è usato per selezionare un insieme di rappresentanti $M \subseteq S^2$ per le orbite dell'azione del gruppo generato da due rotazioni indipendenti. Il risultato è così paradossale che può far nascere dei dubbi sull'assioma della scelta, anche se la maggior parte dei matematici concorda sul fatto che rinunciarvi porterebbe più svantaggi che vantaggi. I teorici degli insiemi hanno tuttavia considerato un'alternativa all'assioma della scelta, chiamato *assioma di determinatezza*, il quale implica tra le altre cose che tutti gli insiemi siano Lebesgue-misurabili e l'enunciato del teorema di Banach-Tarski sia conseguentemente falso. Volendo mantenere i vantaggi di entrambi gli assiomi, si può dimostrare che l'assioma della scelta è compatibile con una forma limitata, ma tuttavia molto potente, di determinatezza.

Accenniamo infine ad un risultato di Dougherty e Foreman che non usa la scelta. Per poterlo enunciare abbiamo bisogno di una definizione. Diciamo che X e Y sono *densamente equiscomponibili* se esistono sottoinsiemi aperti disgiunti X_1, \dots, X_n di X e Y_1, \dots, Y_n di Y tali che $X_1 \cup \dots \cup X_n$ è denso in X , $Y_1 \cup \dots \cup Y_n$ è denso in Y e X_i è isometrico a Y_i per ogni $i = 1, \dots, n$. Si noti che un caso particolare è dato dalle congruenze per tagli, con i vari X_i nel ruolo delle parti interne dei pezzi della congruenza.

Dougherty e Foreman hanno dimostrato in [2], *senza utilizzare l'assioma della scelta*, che due sottoinsiemi aperti e limitati di \mathbb{R}^3 (ad esempio una palla del diametro di una biglia e una del diametro del sole) sono densamente equiscomponibili.

Il risultato è tanto più incredibile se si riflette sul fatto che i pezzi X_i e Y_i nella definizione di densa equiscomponibilità, essendo aperti, sono Lebesgue-misurabili, e anche le loro chiusure lo sono. Per approfondire gli argomenti il lettore potrà consultare [3].

Riferimenti bibliografici

- [1] D. BENKO, *A new approach to Hilbert's third problem*, Amer. Math. Monthly **114** (2007), 665-676
- [2] R. DOUGHERTY, M. FOREMAN *Banach-Tarski decompositions using sets with the property of Baire*, J. Amer. Math. Soc. **7** (1994), 75-124.

- [3] G. TOMKOWICZ, S. WAGON “The Banach-Tarski Paradox”, seconda edizione, Encyclopedia of Mathematics and its Applications, Vol. 163, Cambridge University Press, New York, 2016.

Dipartimento di Matematica
Università di Pisa
Largo Bruno Pontecorvo, 5
56127 PISA – Italy
alessandro.berarducci@unipi.it