

Dehn e Banach-Tarski: opposti paradossi

E. Paolini

26 luglio 2016

Ho ancora bene in mente quando alla scuola elementare abbiamo ricavato le formule per calcolare l'area del triangolo, del trapezio e del rombo. Disegnavamo le forme sul cartoncino e le tagliavamo in modo opportuno in modo che ri assemblando i pezzi si poteva formare un rettangolo. Mi ricordo l'enorme soddisfazione di avere uno strumento che ci permetteva di determinare l'area di qualunque poligono, riducendolo a triangoli e ri assemblando i triangoli in rettangoli.

A quel tempo, mai avrei potuto pensare che queste operazioni, passando dal piano (due dimensioni) allo spazio (tre dimensioni) potessero fallire con due opposti paradossi¹. Se il problema è quello di definire assiomaticamente il concetto di *misura di volume* i teoremi di Dehn, e di Banach-Tarski fissano dei paletti invalicabili e delimitano quindi la regione in cui si può fondare la definizione di volume di un solido.

Il teorema di Dehn afferma infatti che non è sempre possibile *scindere* una piramide in pezzetti poliedrali (le cui parti interne sono a due a due disgiunte) che poi possano essere ricomposti tramite isometrie per formare un parallelepipedo dello stesso volume. Il paradosso di Banach-Tarski afferma al contrario che se ammettiamo qualunque tipo di decomposizione (non solo poliedrale) allora qualunque solido geometrico (cubo, piramide, sfera piena...) può essere decomposto in un numero finito di pezzi (insiemi arbitrari a due a due disgiunti) che ricomposti per isometrie danno origine ad un qualunque altro solido geometrico, indipendentemente dal volume! In particolare (ed è questa la formulazione più comune) è possibile duplicare una palla.

Se, da un lato, questi teoremi sono molto tecnici e rappresentano esempi patologici, dall'altro, se vogliamo comprendere quali siano i percorsi didattici lungo cui è più opportuno sviluppare la teoria geometrica della misura, è importante avere una visione *superiore* e quindi essere a conoscenza dell'esistenza di questi esempi e saper inquadrare i metodi e gli strumenti utilizzati per la loro dimostrazione.

Lo scopo di queste note è quello di presentare questi due importanti teoremi. Vogliamo entrare nel cuore delle dimostrazioni e, senza eccedere in formalismi, cercare di estrarre e riconoscere gli elementi fondamentali che compongono tali dimostrazioni e che sono di per sé interessanti. Vedremo quindi come è possibile tracciare un percorso (con alcune deviazioni) che ci conduce dal paradosso dell'infinito al paradosso di Banach-Tarski, passando dall'assioma della scelta, dall'esempio di Vitali, dai frattali e dal teorema di Dehn.

1 Assiomatica dell'area

Se dobbiamo definire la misura di superficie $\mu(A)$, ovvero l'area, di una regione del piano A , vogliamo rispettare i seguenti assiomi:

¹Il termine *paradosso* viene qui utilizzato per indicare un risultato vero che però contrasta con la nostra intuizione.

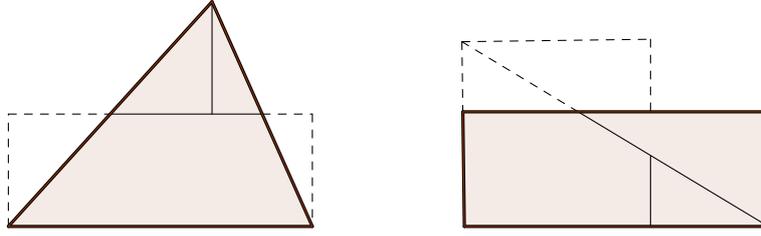


Figura 1: Equiscindibilità di un triangolo a un rettangolo, ed equiscindibilità di un rettangolo a un quadrato.

1. se A è un rettangolo con i lati di lunghezza a e b , allora $\mu(A) = ab$;
2. additività: se A e B sono tra loro disgiunti, allora $\mu(A \cup B) = \mu(A) + \mu(B)$;
3. invarianza per isometrie: se θ è una isometria del piano (rotazione, traslazione, riflessione), allora $\mu(\theta(A)) = \mu(A)$.

Queste proprietà garantiscono l'invarianza della misura μ per *equiscissione* (dall'inglese: *scissor-equivalence*): se un poligono A può essere suddiviso in poligoni disgiunti (si intende che i poligoni possono sovrapporsi solamente sui lati): $A = A_1 \cup \dots \cup A_n$ e se con opportune isometrie θ_i si ottengono dei poliedri $B_i = \theta_i(A_i)$ ancora a due a due disgiunti e poniamo $B = \theta_1(A_1) \cup \dots \cup \theta_n(A_n)$, allora $\mu(A) = \mu(B)$. Questa proprietà discende dagli assiomi, in quanto per additività si ha $\mu(A) = \mu(A_1) + \dots + \mu(A_n)$ e $\mu(B) = \mu(B_1) + \dots + \mu(B_n)$ mentre dall'invarianza per isometrie si ha $\mu(B_i) = \mu(\theta_i(A_i)) = \mu(A_i)$. L'unico dettaglio su cui si potrebbe discutere è il fatto che nella equiscissione si ammette che i poligoni ritagliati possano sovrapporsi sui bordi. Questo può essere giustificato dal fatto che i lati di un poligono sono dei rettangoli degeneri di altezza nulla e quindi, per il primo assioma, avranno area nulla. Dunque l'additività è comunque valida in quanto l'additività sui disgiunti ci dice che in generale (anche se A e B non sono disgiunti) si ha

$$\mu(A) + \mu(B) = \mu(A \setminus B) + \mu(B \setminus A) + 2\mu(A \cap B) = \mu(A \cup B) + \mu(A \cap B)$$

quindi $\mu(A \cap B) = 0$ è sufficiente per avere l'additività.

E' facile utilizzare l'equiscindibilità per trovare una suddivisione di un generico triangolo di base b e altezza h in tre poligoni che ricomposti danno origine ad un rettangolo di base b e altezza $h/2$ (si veda la figura 1). E visto che ogni poligono può essere suddiviso in triangoli, ogni poligono è equiscindibile ad una unione disgiunta di rettangoli (Teorema di Wallace–Bolyai–Gerwien, [3]). Dunque i tre assiomi che abbiamo sopra enunciato ci permettono di definire, in modo univoco, l'area μ di tutti i poligoni del piano.

Si può anche dimostrare che ogni poligono è equiscindibile ad un quadrato, mostrando che ogni rettangolo è equiscindibile ad un quadrato (cosa non del tutto banale, ma si veda la figura 1). Inoltre si può dimostrare che l'equiscindibilità dei poligoni, quando c'è, si può sempre ottenere utilizzando come isometrie solamente le traslazioni (evitando quindi rotazioni e riflessioni). Dunque negli assiomi enunciati si può richiedere solamente una invarianza per traslazione (al posto dell'invarianza per isometria) e si può fissare l'area dei soli quadrati (invece che di tutti i rettangoli) e comunque si avrebbe una definizione univoca di area dei poligoni (si veda ad esempio [3]).

2 Assiomatica del volume

Se la formula per l'area del triangolo è nota ed è comprensibile fin dalla scuola elementare, la formula del volume della piramide (area di base moltiplicata per altezza diviso tre) ci è stata probabilmente propinata alla scuola media inferiore senza una spiegazione completamente convincente. Tale formula può essere dimostrata utilizzando dei procedimenti di approssimazione (gli stessi che si usano per trovare le aree o i volumi di regioni curve). Ma, ci chiediamo, è possibile che non ci sia un argomento elementare, simile a quello utilizzato per i triangoli?

Detto in altri termini: è vero che ogni poliedro dello spazio è equiscindibile ad un parallelepipedo retto? Se questa domanda ci sembra sensata siamo contenti di sapere che anche Gauss se l'era posta e congetturò che la risposta fosse negativa. La stessa domanda è stata poi ripresa nel terzo problema di Hilbert, all'inizio del secolo scorso, ed è stata quindi risolta in maniera negativa da Dehn, allievo dello stesso Hilbert: non è possibile suddividere un tetraedro regolare in poliedri disgiunti che ricomposti diano origine ad un parallelepipedo retto (si vedano in proposito [1] e [5]).

Vedremo nel seguito la dimostrazione di questo sorprendente risultato.

Il teorema di Dehn ci fa comprendere che gli assiomi che abbiamo postulato non sono sufficienti a determinare univocamente la misura di volume. Quello che ci manca è un assioma che ci permetta di calcolare le misure per approssimazione, ad esempio il seguente:

4. monotonia: se $A \subseteq B$ allora $\mu(A) \leq \mu(B)$.

Questo assioma oltre ad essere necessario, nello spazio, per determinare la misura di tutti i poliedri, ci permette (sia nello spazio che nel piano) di determinare la misura di figure curve. L'assioma di monotonia ci permette, ad esempio, di determinare l'area del cerchio approssimando la circonferenza dall'esterno e dall'interno con poligoni e osservando che la differenza tra l'area del poligono esterno e quella del poligono interno può essere resa piccola a piacere. La stessa cosa si può fare per calcolare il volume della sfera o del cono con base qualunque (e quindi anche della piramide). Tale assioma può essere utilizzato per definire la misura di Peano-Jordan e l'integrale di Riemann. Con questi strumenti si può anche dimostrare il principio di Cavalieri il quale, anch'esso, può essere utilizzato per determinare il volume del cono e della sfera.

Abbiamo fin qui visto come il teorema di Dehn mostra la carenza dei primi tre assiomi che abbiamo presentato, mostrando come essi non siano sufficienti ad ottenere una definizione univoca di misura sugli insiemi poliedrali. Da un lato, dunque, questi tre assiomi sembrano troppo deboli.

D'altro canto il risultato di Banach-Tarski ci dice che questi tre assiomi sono invece talmente forti da non essere soddisfacibili! Una possibile formulazione del paradosso di Banach-Tarski ci dice infatti che dati due parallelepipedi qualunque A e B è possibile suddividere A in un numero finito di insiemi a due a due disgiunti: $A = A_1 \cup \dots \cup A_n$ tali che con opportune isometrie θ_i si ottengono degli insiemi $B_i = \theta_i(A_i)$, anch'essi a due a due disgiunti e tali che l'unione dei B_i è il parallelepipedo B . Se allora esistesse una misura invariante per isometrie, si avrebbe $\mu(B_i) = \mu(\theta_i(A_i)) = \mu(A_i)$ e se la misura fosse additiva si avrebbe $\mu(A) = \mu(B)$ dunque si avrebbe che tutti i parallelepipedi dovrebbero avere lo stesso volume, da cui, sempre per l'additività, si ricaverebbe facilmente che tutti i parallelepipedi dovrebbero avere volume nullo, invalidando il primo assioma.

Questo risultato paradossale può essere risolto solamente con l'osservazione che i pezzi utilizzati nella decomposizione di Banach-Tarski sono insiemi decisamente patologici e non costruibili. Si dovrà quindi ammettere che la misura μ non possa

essere definita su qualunque insieme, ma solo su una categoria di insiemi *buoni* (gli insiemi *misurabili*, appunto). E quindi l'additività sarà richiesta solamente sulle suddivisioni di insiemi misurabili tramite pezzi misurabili.

Nel seguito vedremo la dimostrazione di questo risultato di Banach-Tarski che è assolutamente non banale e farà pesante uso del paradosso dell'infinito e dell'assioma della scelta, oltre che della struttura non abeliana del gruppo di rotazioni nello spazio (raccomandiamo in proposito [7] e [8]).

Ci limitiamo qui ad osservare che il paradosso di Banach-Tarski vale nello spazio ma non nel piano (l'analogia con il teorema di Dehn è evidente anche se apparentemente i due risultati non sono tra loro connessi). In effetti esistono misure non nulle, additive, invarianti per isometrie e definite su tutti i sottoinsiemi del piano (misure di Banach).

Questa discrepanza tra il caso planare e il caso spaziale (il teorema di Banach-Tarski vale in qualunque dimensione maggiore di due) viene a cadere se al posto della *additività* (finita) si richiama la σ -*additività* (ovvero additività su unioni numerabili). L'estensione dell'assioma di additività sulle unioni infinite (ma comunque numerabili) è importante per poter definire la misura (di Lebesgue) su una classe più ampia di insiemi che sia stabile per approssimazione. Ma se si richiede tale proprietà aggiuntiva è molto più facile dimostrare che la misura non può essere definita su tutti gli insiemi. L'esistenza di insiemi non misurabili per misure σ -additive si ottiene in qualunque dimensione ed è dovuta a Vitali. Vedremo nel seguito anche questo esempio che, pur essendo molto più semplice, ha una costruzione che viene ripresa anche nella dimostrazione di Banach-Tarski.

Nel seguito cercheremo di fornire una dimostrazione dei teoremi di Dehn e di Banach-Tarski. Per introdurre queste dimostrazioni in maniera più "morbida" e per mettere in evidenza le similarità con altre problematiche più elementari, esploreremo preventivamente alcuni altri risultati propedeutici.

3 Il paradosso dell'infinito: l'hotel Hilbert

Nel momento in cui Cantor ha introdotto la teoria degli insiemi e quindi la possibilità di trattare in maniera coerente collezioni infinite di oggetti, ci si è posto il problema di adeguare la nostra intuizione a questi enti, che difficilmente hanno un riscontro nella realtà.

L'hotel Hilbert è un normale albergo, in cui ogni stanza è etichettata con un numero naturale. Stanze diverse hanno numeri diversi (iniettività dell'etichettatura). L'unica particolarità di questo albergo è che questa corrispondenza è *biunivoca* ovvero che dato qualunque numero naturale, esiste una stanza con quel numero. In altre parole ci sono infinite stanze e per la precisione: una quantità numerabile di stanze. Sappiamo come la storia va avanti. Un cliente si presenta alla *reception* in un momento in cui tutte le stanze dell'hotel sono occupate. Non c'è spazio per un nuovo cliente. . . o forse c'è? Il direttore dell'hotel chiede gentilmente ad ogni cliente di cambiare stanza, in particolare il cliente che si trova nella stanza n deve andare nella stanza $n + 1$. In questo modo ogni cliente ha la sua stanza, e la stanza numero 1 si è liberata.²

L'hotel Hilbert è uno dei paradossi dell'infinito o, meglio, è *il* paradosso dell'infinito, visto che è sostanzialmente equivalente ad una delle possibili definizioni di insieme infinito: un insieme A è infinito se esiste una funzione $\sigma: A \rightarrow A$ iniettiva ma non suriettiva.

Nell'esempio dell'hotel Hilbert abbiamo utilizzato la funzione $\sigma: \mathbb{N} \rightarrow \mathbb{N}$ definita da $\sigma(n) = n + 1$. L'esistenza di una tale funzione (insieme alla richiesta che

²Decidiamo qui, per analogia con i numeri delle stanze d'albergo, che il primo numero naturale sia 1.

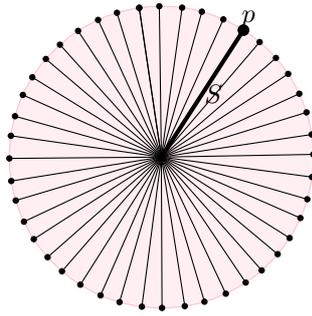


Figura 2: La *ventola* che tramite una rotazione permette di assorbire o generare un punto.

sia iniettiva e non suriettiva) è insita negli assiomi di Peano dei numeri naturali (numeri diversi hanno diversi successori, l'uno non è successore di alcun numero). Se immergiamo i numeri naturali nella retta reale, osserviamo inoltre che questa mappa σ è una isometria, più precisamente una traslazione. Abbiamo quindi la nostra prima decomposizione di *assorbimento*: l'insieme \mathbb{N} , dei numeri naturali, può essere decomposto in due pezzi: $A = \{1\}$ e $B = \mathbb{N} \setminus \{1\} = \sigma(\mathbb{N})$. Tolto A rimane quindi un insieme isometrico all'originale. Questo significa che se abbiamo a che fare con insiemi infiniti, è possibile generare punti *dal nulla* semplicemente spostando l'insieme.

Si potrebbe pensare che questa decomposizione di \mathbb{N} dipenda dal fatto che abbiamo un insieme illimitato. Ma non è questo il caso. Infatti nel piano possiamo costruire un insieme limitato che abbia le stesse caratteristiche paradossali di \mathbb{N} . Sarà sufficiente avvolgere \mathbb{N} sulla circonferenza unitaria... visto che π è irrazionale, i punti di \mathbb{N} non andranno mai a sovrapporsi e l'insieme risultante avrà una analoga proprietà di assorbimento: se togliamo il primo punto e ruotiamo l'insieme rimanente di un radiante, otteniamo nuovamente l'insieme originario. Una sorta di ventola infinita.

Questo esempio è piuttosto semplice, ma ci sarà utile (visto che poi faremo una costruzione simile per le rotazioni dello spazio) rivisitarlo con una notazione algebrica più astratta. Fissata una rotazione θ del piano, possiamo considerare il gruppo $G = \langle \theta \rangle$ delle rotazioni generate da θ . Si intende cioè che G è il gruppo di isometrie contenente l'identità $e = \theta^0$, la rotazione $\theta = \theta^1$ la sua inversa $\bar{\theta} = \theta^{-1}$ e poi ogni rotazione θ^n che si ottiene applicando n volte θ , e ogni rotazione θ^{-n} che si ottiene applicando n volte la rotazione inversa $\bar{\theta}$. Dunque $G = \{\theta^n : n \in \mathbb{Z}\}$.

Se θ è una rotazione di un angolo commensurabile con l'angolo giro, allora esisterà un n tale che $\theta^n = e$. In tal caso il gruppo G sarà finito e isomorfo a \mathbb{Z}_n . Se invece θ è una rotazione di un angolo incommensurabile con l'angolo giro il gruppo G sarà isomorfo a \mathbb{Z} e avrà, in particolare, infiniti elementi. Chiamiamo $G_\theta = \{\theta^n : n \in \mathbb{N} \cup \{0\}\}$ il semigruppato generato da θ (senza usare mai la rotazione inversa). Fissato un punto p che non sia il centro della rotazione θ , possiamo allora costruire l'insieme $A_\theta = G_\theta(\{p\}) = \{\theta^n(p) : n \in \mathbb{N}\}$. Chiamiamo tale insieme *orbita* di p : esso è composto da infiniti punti che si addensano su una circonferenza. Si ha allora la seguente decomposizione:

$$A_\theta = \{p\} \cup \theta(A_\theta)$$

cioè è possibile togliere un punto da A_θ in modo che l'insieme rimanente risulti isometrico all'intero A_θ .

Se la stessa costruzione la facciamo a partire da un raggio S invece che da un punto p l'orbita $G_\theta(S)$ sarà un insieme denso in un disco, e avrà la caratteristica di

poter assorbire un intero segmento:

$$G_\theta(S) = S \cup \theta(G_\theta(S)).$$

L'assorbimento verrà utilizzato nel teorema di Banach-Tarski per ricomprendere nella decomposizione quei piccoli *trucioli* che possono avanzare nel processo centrale della dimostrazione. Per prendere confidenza con questo metodo consideriamo il seguente problema.

Se due insiemi poligonali sono equiscindibili sono anche equidecomponibili? Ricordiamo che l'equiscindibilità richiede una decomposizione in poligoni che abbiano a due a due parte interna disgiunta. Nella equidecomposizione, invece, gli insiemi componenti posso essere qualunque, ma devono avere a due a due intersezione vuota.

I *pezzi* dei poligoni che compongono una scissione possono essere suddivisi in parte interna e perimetro. Il perimetro sarà formato da un numero finito di segmenti. Scelto un $\varepsilon > 0$ questi segmenti potranno essere suddivisi ulteriormente in un numero finito N di segmenti *semiaperti* ognuno di lunghezza inferiore a ε . L'idea è quindi quella di considerare un cerchio di raggio ε tutto contenuto in uno dei poligoni e in questo cerchio andare a considerare N angoli $\theta_1, \dots, \theta_N$ a due a due incommensurabili tra loro (e incommensurabili con l'angolo giro). Andando quindi a posizionare opportunamente gli N segmenti come raggi di questo cerchio, e andando a ruotare ogni segmento tramite i multipli interi del corrispondente angolo θ_k , si otterranno N ventole sovrapposte ma disgiunte, ognuna delle quali, ruotando, potrà generare o assorbire uno degli N segmenti.

Dunque dai poligoni che determinano l'equiscissione possono essere scorporati i lati che si sovrappongono, inoltre andranno individuate le *ventole* e anch'esse andranno scorporate. Quella che si ottiene è una decomposizione fatta da poligoni semiaperti, segmenti e un certo numero di ventole estratte da uno dei poligoni. Ricomponendo i poligoni ci saranno dei segmenti che mancano o avanzano, e questi potranno essere recuperati con opportune rotazioni delle ventole.

4 Il teorema di Cantor-Bernstein

E' possibile suddividere un quadrato in un numero finito di pezzi disgiunti che possano essere ricomposti mediante isometrie e *omotetie* in modo da ottenere un cerchio?

Rispetto alla equidecomponibilità, visto che abbiamo aggiunto la possibilità di utilizzare le omotetie (ovvero riscalamenti dei singoli pezzi), il fatto che il cerchio e il quadrato abbiano o meno la stessa area non è rilevante in questo problema. La difficoltà sembra invece essere la possibilità di decomporre un oggetto curvo e far combaciare i pezzi perfettamente in modo da ottenere un oggetto squadrato. Osserviamo anche, come abbiamo già anticipato, che il teorema di Banach-Tarski non è valido nel piano e quindi neanche esso risponde a questa domanda.

La risposta, affermativa, a questo problema è in realtà molto semplice. A meno di omotetie possiamo supporre che il cerchio C sia contenuto all'interno del quadrato Q . Consideriamo quindi una omotetia f che manda il quadrato all'interno del cerchio: $f(Q) \subseteq C$. Chiamiamo ora $B_1 = Q \setminus C$ la buccia esterna, e costruiamo una successione di bucce $B_2 = f(B_1)$, $B_3 = f(B_2)$, ... in modo da decomporre il quadrato come fosse una cipolla. L'unione di tutte le bucce è un insieme $B = \bigcup_{k \in \mathbb{N}} B_k$ che ha la stessa struttura dei numeri naturali (si veda la figura 3 dove B è la parte grigia) e ci permette quindi di attuare la procedura di assorbimento:

$$f(B) = \bigcup_{k \in \mathbb{N}} f(B_k) = \bigcup_{k \in \mathbb{N}} B_{k+1} = B \setminus B_1.$$

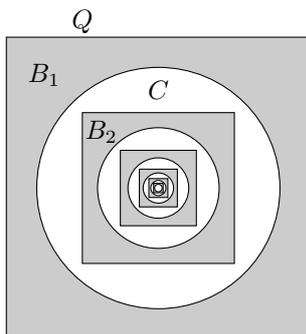


Figura 3: La suddivisione utilizzata per trasformare un quadrato Q in un cerchio C mediante una decomposizione e una omotetia. L'insieme delle bucce B è ombreggiato.

Dunque, posto $A = Q \setminus B$ abbiamo la decomposizione cercata:

$$Q = A \cup B, \quad C = A \cup f(B).$$

Questa semplice idea ci permette di dimostrare il teorema di Cantor-Bernstein cioè il teorema che ci garantisce che la cardinalità degli insiemi è una relazione d'ordine totale. Due insiemi A e B si dicono avere la stessa *cardinalità* o essere *equipotenti* se esiste una corrispondenza biunivoca tra di essi. In tal caso scriveremo $|A| = |B|$. Se un insieme A ha la stessa cardinalità di un sottoinsieme di B , allora scriveremo $|A| \leq |B|$.

Il teorema di Cantor-Bernstein ci dice che se $|Q| \leq |C|$ e, contemporaneamente, $|C| \leq |Q|$ allora $|Q| = |C|$ (totale ordinamento degli insiemi rispetto alla cardinalità). Questo risultato si usa continuamente nelle dimostrazioni di equipotenza di insiemi, in particolare nei risultati di Cantor sulla cardinalità degli insiemi numerici: $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| < |\mathbb{R}| = |C|$.

Nell'enunciato del teorema di Cantor-Bernstein abbiamo denotato con Q e C i due insiemi perché, in effetti, la stessa dimostrazione che abbiamo fatto per quadrato e cerchio e per le omotetie, si ripete sostanzialmente invariata in questo contesto più generale.

Per ipotesi abbiamo $|Q| \leq |C|$ e $|C| \leq |Q|$. La seconda relazione ci dice che esiste una funzione iniettiva $g: C \rightarrow Q$. Consideriamo $C' = g(C)$ cosicché $C' \subseteq Q$. Sarà allora sufficiente dimostrare che $|C'| = |Q|$ in quanto la funzione $g: C \rightarrow C'$ è una bigezione e dunque $|C| = |C'|$ (la proprietà transitiva dell'equipotenza è banale, in quanto la composizione di funzioni biunivoche è anch'essa biunivoca).

Senza perdita di generalità possiamo quindi assumere che si abbia $C = g(C)$ cioè $C \subseteq Q$. L'ipotesi rimanente è dunque l'esistenza di una funzione iniettiva $f: Q \rightarrow C$. La dimostrazione procede quindi con la decomposizione a *cipolla* rappresentata in figura 3. Il fatto che f sia una generica funzione iniettiva, e non una omotetia (concetto che non ha senso per insiemi astratti), non è in effetti rilevante e la dimostrazione si conclude facendo la stessa partizione già vista:

$$Q = A \cup B, \quad C = A \cup f(B)$$

e definendo la funzione biunivoca $h: Q \rightarrow C$

$$h(x) = \begin{cases} f(x) & \text{se } x \in B \\ x & \text{se } x \in A. \end{cases} \quad (1)$$

Questo conclude il teorema di Cantor-Bernstein per quanto riguarda la cardinalità. Vogliamo però ora estendere questo risultato (e questo ci servirà in seguito)

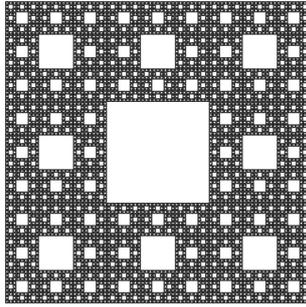


Figura 4: Il tappeto di Sierpinski.

alla equidecomposizione. Introduciamo la notazione $A \sim B$ per indicare che A e B sono equidecomponibili. E usiamo la notazione $A \preceq B$ se A è equidecomponibile ad un sottoinsieme di B . Ora se $C \subset Q$ e $Q \preceq C$ si ha dunque una partizione finita di $Q = Q_1 \cup \dots \cup Q_N$ e una mappa $f: Q \rightarrow C$ che è una isometria se ristretta ad ognuno dei pezzi Q_k . In particolare tale mappa è iniettiva. Si può dunque ripetere la dimostrazione appena vista, e costruire l'insieme di bucce B osservando che la mappa biunivoca $h: Q \rightarrow C$ risultante, si scrive come in (1).

Vediamo allora che a patto di dividere in due ognuno dei pezzi Q_k : $Q'_k = Q_k \cap B$, $Q''_k = Q_k \setminus B$, si ottiene una nuova partizione di Q con un numero al più doppio di pezzi sui quali la mappa h risulta essere una isometria (perché sia f che l'identità sono isometrie su ogni Q_k).

Dunque abbiamo dimostrato il seguente risultato: se A è equidecomponibile ad un sottoinsieme di B (cioè $A \preceq B$) e B è equidecomponibile ad un sottoinsieme di A (cioè $B \preceq A$), allora A e B sono equidecomponibili (cioè $A \sim B$).

5 Frattali autosimili e dimensione

Diremo che una similitudine S ha rapporto λ se la distanza dei punti varia con coefficiente λ :

$$d(S(x), S(y)) = \lambda d(x, y).$$

Se $\lambda > 1$ si tratta di una dilatazione, se $0 < \lambda < 1$ si tratta di una contrazione, se $\lambda = 1$ si tratta di una isometria. Ci aspettiamo che una similitudine S_λ di rapporto λ vari la misura μ di un insieme tramite una legge di questa forma:

$$\mu(S_\lambda(A)) = \lambda^d \mu(A). \quad (2)$$

L'esponente d è chiamato *dimensione* della misura. Per una lunghezza si avrà $d = 1$, per un'area $d = 2$, per un volume $d = 3$. Questa proprietà risulta chiara se osserviamo, ad esempio, che un segmento di lunghezza 2 si decompone in $2^1 = 2$ segmenti di lunghezza 1. Un quadrato di lato 2 si decompone invece in $2^2 = 4$ quadrati di lato 1 e un cubo di spigolo 2 in $2^3 = 8$ cubi di lato 1. Dunque la relazione (2) discende, in questi casi, dall'additività della misura presa in considerazione.

Consideriamo ora l'insieme T (tappeto di Sierpinski) in figura 4. Questo insieme, per costruzione, è formato da 8 copie riscalate di se stesso. Ognuna delle otto copie si ottiene dall'originale applicando una traslazione e poi una omotetia di coefficiente $\lambda = 1/3$. Dunque se μ è una misura di dimensione d si ottiene:

$$\mu(T) = 8\mu(S_{\frac{1}{3}}(T)) = \frac{8}{3^d} \mu(T). \quad (3)$$

Questa equazione è certamente soddisfatta se $\mu(T) = 0$ oppure $\mu(T) = +\infty$. In effetti, come vedremo tra breve, ha senso affermare che la *lunghezza* di questo

insieme è infinita mentre la sua *area* è nulla. Osserviamo però che sarebbe possibile soddisfare tale equazione anche con $0 < \mu(T) < +\infty$ se si avesse $8/3^d = 1$ e cioè se $d = \log_3 8 = 1,892\dots$. Tale valore di d può essere, a buona ragione, assunto come valore frazionario della dimensione dell'insieme T che quindi risulta essere una via di mezzo tra un insieme unidimensionale ($d = 1$ come succede in una curva) ed un insieme bidimensionale ($d = 2$ come accade per le superfici) e per questo motivo ci si aspetta che T abbia lunghezza infinita e area nulla. Misure di dimensione d arbitraria esistono effettivamente, ad esempio le *misure di Hausdorff* hanno questa proprietà e possono essere quindi utilizzate per determinare la dimensione e misurare l'estensione di insiemi frattali (si veda [2]).

Nella dimostrazione del teorema di Banach-Tarski utilizzeremo il grafo frattale di figura 5. Vedremo come anche in quel caso saranno importanti le proprietà di autosimilarità di quella figura per ottenere una riproduzione dell'intero insieme tramite una sua parte propria.

6 L'assioma della scelta

A partire dall'esempio di Vitali (che vedremo tra breve) e sia nel teorema di Dehn che nel teorema di Banach-Tarski, si fa uso dell'assioma della scelta. La teoria degli insiemi, introdotta da Cantor, viene usualmente formalizzata tramite gli assiomi di Zermelo-Fraenkel (ZF). Tra questi assiomi troviamo l'assioma di infinito, che è sostanzialmente la proprietà evidenziata nell'esempio dell'hotel Hilbert. Questo assioma è necessario per garantire l'esistenza di insiemi infiniti, in particolare garantisce la possibilità di definire l'insieme dei numeri naturali dal quale vengono poi costruiti gli altri insiemi utilizzati nella matematica di base.

Abbiamo già messo in evidenza il fatto che l'assioma di infinito di per sé può essere considerato paradossale. E, in effetti, i teoremi di Dehn e di Banach-Tarski possono essere considerati un modo estremo di evidenziare il lato paradossale di questo assioma.

L'altro ingrediente fondamentale per la costruzione degli esempi di Dehn e Banach-Tarski è però un assioma aggiuntivo al sistema ZF: l'assioma della scelta (AC dall'inglese *Axiom of Choice* per cui il sistema assiomatico ZF esteso con AC prende il nome di ZFC). Un'ampia trattazione di questo delicatissimo enunciato si trova, ad esempio, in [4].

L'assioma della scelta riguarda la possibilità di poter scegliere, data una famiglia arbitraria di insiemi, un elemento da ogni insieme (quella che chiameremo nel seguito una *sezione*) anche in assenza di un criterio esplicito per fare tale scelta. Questa possibilità è del tutto naturale, ed è ovvia se si pensa agli insiemi finiti. E' invece sorprendente il fatto che tale possibilità non è garantita dagli altri assiomi della teoria degli insiemi, e dunque risulta che il sistema ZFC è effettivamente una estensione di ZF (Cohen, 1963).

L'assioma della scelta viene spesso utilizzato per costruire esempi patologici: insiemi non misurabili (l'esempio di Vitali), le funzioni additive ma non lineari, la decomposizione di Dehn e la decomposizione di Banach-Tarski. Se questi risultati *paradossali* non ci dovessero piacere, però, non sarebbe sufficiente rimuovere l'assioma della scelta. Senza l'assioma della scelta, infatti, l'esistenza di tali esempi patologici non verrebbe negata ma diverrebbe una proposizione indecidibile. La rimozione dell'assioma della scelta non può che diminuire il numero di teoremi dimostrabili e quindi non ci permetterebbe di escludere, ad esempio, l'esistenza di insiemi non misurabili o di decomposizioni paradossali. Per avere teoremi che escludano l'esistenza di questi esempi patologici sarebbe necessario aggiungere degli assiomi alternativi ad AC e con esso incompatibili (ad esempio Solovay [6] ha costruito un modello di ZF in cui AC non vale ed ogni insieme è misurabile).

7 L'esempio di Vitali

L'esempio di Vitali viene introdotto in tutti i corsi di teoria della misura in quanto mostra l'impossibilità di definire una funzione d'insieme definita su tutti i sottoinsiemi limitati di \mathbb{R}^n che sia: non identicamente nulla, invariante per traslazioni e σ -additiva (cioè additiva su unioni numerabili).

Ricordiamo che μ si dice essere σ -additiva se vale

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} \mu(A_n)$$

su ogni famiglia A_n di insiemi a due a due disgiunti.

Nell'esempio di Vitali si costruisce una unione numerabile di insiemi su cui non può valere la σ -additività. Nel teorema di Banach-Tarski (molto più complicato da dimostrare) si viola anche la semplice additività finita. Osserviamo qui che oltre ad essere più semplice da dimostrare, l'esempio di Vitali, a differenza di Banach-Tarski, è valido anche in dimensione bassa: sia nel piano che sulla retta. Vedremo qui la costruzione fatta nella retta reale.

Consideriamo l'intervallo semiaperto $I = [0, 1)$ in \mathbb{R} . Dato un punto $x \in I$ diremo che $y \in I$ è \mathbb{Q} -raggiungibile da x se $y - x \in \mathbb{Q}$ (cioè y si ottiene da x aggiungendo un numero razionale). Tutti i punti di I possono essere suddivisi in classi di equivalenza in base alla \mathbb{Q} -raggiungibilità. L'assioma della scelta ci garantisce l'esistenza di una sezione Δ ovvero di un insieme contenente un unico rappresentante di ogni classe. Detto in altre parole per ogni $y \in I$ esiste un unico $x \in \Delta$ tale che y sia \mathbb{Q} -raggiungibile da x .

Viceversa fissato $q \in \mathbb{Q} \cap I$ possiamo considerare l'insieme $\Delta + q$ cioè l'insieme Δ traslato verso destra di una distanza q . Chiamiamo $\Delta_q^+ = (\Delta + q) \cap I$ la parte di Δ che dopo essere stata translata rimane dentro l'intervallo di riferimento I . E chiamiamo $\Delta_q^- = (\Delta + q - 1) \cap I$ che corrisponde a prendere la parte sporgente e riportarla dentro I tramite una traslazione a sinistra di una unità. Chiaramente

$$I = \bigcup_{q \in \mathbb{Q} \cap I} \Delta_q^+ \cup \Delta_q^- \quad (4)$$

in quanto per ogni punto $y \in I$ esiste un unico $x \in \Delta$ tale che $q = y - x \in \mathbb{Q}$: se $y \geq x$ allora $q \in \mathbb{Q} \cap I$ e quindi $y \in \Delta_q^+$ altrimenti $q + 1 \in \mathbb{Q} \cap I$ e quindi $y \in \Delta_q^-$.

La decomposizione (4) presenta l'intervallo I come unione disgiunta numerabile di copie traslate di Δ . Se μ fosse una misura σ -additiva si avrebbe dunque:

$$\mu(I) = \sum_{q \in \mathbb{Q} \cap I} \mu(\Delta_q^+) + \mu(\Delta_q^-) = \sum_{q \in \mathbb{Q} \cap I} \mu(\Delta).$$

Ora se $\mu(\Delta) = 0$ si avrebbe $\mu(I) = 0$, mentre se $\mu(\Delta) > 0$ si avrebbe $\mu(I) = +\infty$. Entrambi i casi escludono la possibilità di avere una misura μ non banale, σ -additiva e definita su tutti i sottoinsiemi limitati della retta reale.

Dunque l'insieme Δ che abbiamo appena costruito dovrà necessariamente essere non misurabile.

Questo esempio, costruito nel caso unidimensionale, può essere esteso facilmente a qualunque dimensione e dunque vale nel piano per la misura d'area e nello spazio per la misura di volume.

Nel piano si può inoltre costruire un interessante variante di questo esempio. Infatti se l'intervallo I viene avvolto attorno ad una circonferenza di lunghezza 1 in modo tale da farne coincidere gli estremi, e se consideriamo l'insieme Σ dei raggi aperti che insistono sui punti della circonferenza corrispondenti all'insieme non

misurabile Δ , si otterrebbe una decomposizione del cerchio C (aperto e senza centro) tramite una unione disgiunta numerabile di copie ruotate dello stesso insieme:

$$C = \bigcup_{q \in \mathbb{Q} \cap (0,1)} R_{2\pi q}(\Sigma) \quad (5)$$

(con R_θ abbiamo indicato la rotazione di un angolo pari a θ radianti).

Abbiamo quindi verificato che non può esistere una funzione d'insieme definita su tutti i sottoinsiemi limitati di \mathbb{R}^n che sia: non identicamente nulla, invariante per traslazioni e σ -additiva (cioè additiva su unioni numerabili). L'esempio di Vitali ci costringe a limitare le nostre richieste, e quindi ci mette di fronte a due alternative (entrambe valide): la prima è quella di non definire la misura di tutti gli insiemi limitati, ma solo di una parte di essi, la seconda è quella di rimpiazzare l'additività con una più debole *sub-additività*. In entrambi i casi si definisce la classe degli insiemi *misurabili* che corrisponde alla famiglia di quegli insiemi sui quali l'additività risulta soddisfatta.

8 Funzioni additive ma non lineari

Una funzione $\phi: \mathbb{R} \rightarrow \mathbb{R}$ si dice essere *additiva* se soddisfa la proprietà:

$$\phi(x + y) = \phi(x) + \phi(y).$$

Esempi molto semplici di funzioni additive sono dati dalle funzioni *lineari* ovvero le funzioni della forma $\phi(x) = \lambda x$. Ci chiediamo: esistono funzioni additive che non siano lineari?

La domanda è più complessa di quanto non possa sembrare a prima vista. Se fissiamo un valore di una funzione additiva, ad esempio fissiamo il valore $\phi(x_1)$ su un numero $x_1 \neq 0$, possiamo facilmente scoprire che $\phi(nx_1) = n\phi(x_1)$ e che $\phi(x_1/n) = \phi(x_1)/n$ per ogni n intero. Si ottiene quindi che $\phi(qx_1) = q\phi(x_1)$ per ogni $q \in \mathbb{Q}$. Dunque se fissiamo ϕ su un numero x_1 , la funzione risulta di conseguenza fissata su tutto $\mathbb{Q}x_1$ ovvero su tutti i numeri commensurabili con x_1 . Allora possiamo scegliere un $x_2 \notin \mathbb{Q}x_1$ e fissare il valore su x_2 . A questo punto la funzione risulta determinata anche su $\mathbb{Q}x_2$ ma non solo, risulta univocamente determinata su tutto $\mathbb{Q}x_1 + \mathbb{Q}x_2$ ovvero sulle combinazioni lineari a coefficienti razionali di x_1 e x_2 . Infatti si dovrà avere

$$\phi(q_1x_1 + q_2x_2) = \phi(q_1x_1) + \phi(q_2x_2) = q_1\phi(x_1) + q_2\phi(x_2).$$

Comprendiamo allora che le funzioni additive altro non sono che funzioni lineari dello spazio vettoriale $\mathbb{R}_{\mathbb{Q}}$ in sé, dove $\mathbb{R}_{\mathbb{Q}}$ è lo spazio vettoriale \mathbb{R} considerato però sul campo \mathbb{Q} invece che su \mathbb{R} stesso. Le funzioni lineari di questo spazio possono essere chiamate \mathbb{Q} -lineari, in quanto sono lineari solo sulle combinazioni fatte con coefficienti in \mathbb{Q} . Lo spazio $\mathbb{R}_{\mathbb{Q}}$ è uno spazio vettoriale di dimensione infinita. L'esistenza di una base algebrica di questo spazio (come di ogni spazio vettoriale) è garantita dal Lemma di Zorn che è una conseguenza non banale dell'assioma della scelta (e risulta essere ad esso equivalente). Non vogliamo entrare nei dettagli della questione, se non accennare al fatto che il Lemma di Zorn si applica prendendo in considerazione le due condizioni utilizzate nell'induzione transfinita.

La prima condizione è che se B è un insieme indipendente di vettori ma non è una base, allora a B si può aggiungere un vettore (nel nostro caso: un numero reale) che non è combinazione finita a coefficienti nel campo (nel nostro caso \mathbb{Q}) di elementi di B . In questo modo dato un insieme indipendente B si può esibire un insieme indipendente B' con un elemento in più. Aggiungendo un elemento alla

volta non si potrà però andare molto lontano. Se partiamo da un insieme finito (o anche numerabile) otterremo sempre insiemi finiti (o numerabili). Sappiamo invece che una eventuale base di $\mathbb{R}_{\mathbb{Q}}$ dovrà avere una quantità infinita, più che numerabile, di elementi, in quanto le combinazioni finite a coefficienti in \mathbb{Q} di elementi estratti da un insieme numerabile sono una quantità numerabile e non possono quindi ricoprire l'intero insieme \mathbb{R} che è più che numerabile.

La seconda operazione necessaria per attuare l'induzione transfinita è quella che richiede di poter estendere una qualunque *catena* di insiemi B_i che siano uno contenuto nell'altro (l'indice i varierà in un generico insieme totalmente ordinato). Questa estensione è in realtà molto semplice perché se gli insiemi B_i sono contenuti uno nell'altro è sufficiente farne l'unione per ottenerne una estensione. Questa operazione ci permette di attraversare i cosiddetti *ordinali limite* (ad esempio un insieme numerabile è unione di una successione crescente di insiemi finiti) fino a garantire che l'insieme indipendente può essere esteso in modo *massimale* cioè fino ad arrivare ad un insieme B che non ammette ulteriori estensioni. Tale insieme, dunque, sarà una base del nostro spazio vettoriale.

Per riassumere diciamo che l'assioma della scelta ci permette, insieme all'induzione transfinita, di dimostrare il lemma di Zorn. Il lemma di Zorn garantisce la possibilità di estendere qualunque insieme indipendente di vettori di uno spazio vettoriale ad una base B dello spazio stesso. Una funzione additiva ϕ non è altro che una funzione \mathbb{Q} -lineare e quindi è univocamente determinata una volta che se ne scelgano i valori su una qualunque base B di $\mathbb{R}_{\mathbb{Q}}$. In particolare se si scelgono due elementi qualunque $x_1, x_2 \in B$ e due valori qualunque $\lambda_1, \lambda_2 \in \mathbb{R}$ esiste una funzione ϕ additiva e tale che $\phi(x_1) = \lambda_1$ e $\phi(x_2) = \lambda_2$. E se scegliamo λ_1 e λ_2 in modo tale che $\lambda_2/x_2 \neq \lambda_1/x_1$ quella che abbiamo trovato è una funzione additiva ma non \mathbb{R} -lineare.

Le funzioni additive ma non lineari sono decisamente patologiche. Visto che sui numeri razionali tali funzioni sono lineari, non potranno essere continue (altrimenti sarebbero lineari su tutto \mathbb{R} visto che \mathbb{Q} è denso in \mathbb{R}). Certamente sono funzioni discontinue in ogni punto, ma non solo: con qualche semplice osservazione si può comprendere che tali funzioni sono ovunque illimitate, cioè sono illimitate su qualunque intervallo non vuoto.

L'esistenza di tali funzioni patologiche deve essere tenuto in considerazione ogni volta che si parla di additività. Ad esempio quando si introduce la funzione esponenziale, non è sufficiente richiedere che $\exp(x + y) = \exp(x)\exp(y)$ in quanto $\exp \phi(x)$ ha tale proprietà qualunque sia ϕ una funzione additiva. Per avere una definizione assiomatica univoca dell'esponenziale tramite additività, sarà quindi necessario richiedere che esso sia quantomeno una funzione continua oppure monotona oppure almeno localmente limitata. Capita, a volte, di vedere definita la funzione esponenziale senza tener conto di queste considerazioni.

Se ce ne fosse bisogno ricordiamo che l'esistenza di queste funzioni additive ma non lineari richiede l'assioma della scelta. Significa che con l'assioma della scelta è possibile dimostrare che tali funzioni esistono ma non è possibile costruirle o esibirle esplicitamente. D'altra parte, come abbiamo già osservato, se non ci fosse l'assioma della scelta non potremmo comunque dimostrare che tutte le funzioni additive sono lineari e il nostro problema iniziale rimarrebbe senza soluzione: una questione indecidibile.

9 Il teorema di Dehn

Siamo ora pronti per dimostrare che esistono poliedri dello stesso volume che non sono equiscindibili. Per fare ciò costruiremo un *invariante*, ovvero una funzione che associa un numero ad ogni poliedro e tale che se due poliedri sono equiscindibili

la funzione restituisce lo stesso numero. Fatto questo verificheremo che esistono due poliedri con lo stesso volume su cui la funzione assume valori diversi: questo significa che tali poliedri non potranno essere equiscindibili.

L'idea è molto semplice, se P è un poliedro definiamo $d(P)$ nel modo seguente:

$$d(P) = \sum_{s \text{ spigolo di } P} \ell(s)\phi(\alpha(s)) \quad (6)$$

dove $\ell(s)$ è la lunghezza dello spigolo s , $\alpha(s)$ è l'ampiezza in radianti dell'angolo tra le due facce di P che si incontrano in s e ϕ è una qualunque funzione additiva tale che $\phi(\pi/2) = 0$.

Verifichiamo che d è invariante per equiscindibilità. Se il poliedro P viene suddiviso nei poliedri P_1, \dots, P_n vogliamo innanzitutto verificare che

$$d(P) = d(P_1) + \dots + d(P_n). \quad (7)$$

Se consideriamo un qualunque spigolo s di P , nella suddivisione in poligoni, tale spigolo sarà stato suddiviso in n parti s_1, \dots, s_n . Su ognuna di queste parti s_k si appoggeranno un certo numero di poliedri della decomposizione. La somma degli angoli di questi poliedri dovrà dare l'angolo totale $\alpha(s)$. Si avrà dunque, sfruttando la linearità rispetto a $\ell(s)$ e l'additività di ϕ :

$$\ell(s)\phi(\alpha(s)) = \sum_{k=1}^n \ell(s_k)\phi(\alpha(s)) = \sum_{\sigma \subseteq s \text{ spigolo di qualche } P_i} \ell(\sigma)\phi(\alpha(\sigma)).$$

Dunque la formula (7) sarebbe valida se al lato destro si facessero le somme solamente sugli spigoli σ dei poliedri P_i che sono contenuti negli spigoli di S . Per completare la dimostrazione occorre quindi verificare che la somma su tutti gli altri spigoli dei poliedri P_i sia nulla.

Se σ è uno spigolo di qualche P_i che non è contenuto in uno spigolo di P abbiamo due possibilità: o σ è contenuto in una faccia di P , oppure σ è interno a P . In entrambi i casi si può ripetere il ragionamento fatto sopra, considerando tutti i poliedri che si appoggiano sullo stesso spigolo σ , la somma degli angoli determinati da σ su tali poliedri dovrà essere $\alpha = \pi$ se lo spigolo σ si trova all'interno di una faccia di P oppure $\alpha = 2\pi$ se lo spigolo è interno al poliedro P . In entrambi i casi il contributo complessivo dato da questo segmento nella somma al lato destro di (7) è $\ell(\sigma)\phi(\alpha)$ e in entrambi i casi $\phi(\alpha) = 0$ in quanto α è multiplo di $\pi/2$ e abbiamo scelto ϕ in modo tale che $\phi(\pi/2) = 0$.

Dunque, riassumendo, posto che si scelga una funzione additiva ϕ tale che $\phi(\pi/2) = 0$, la quantità $d(P)$ risulta essere additiva su una qualunque suddivisione di P . Ora è anche chiaro che $d(P)$ è invariante per isometrie (angoli e lunghezze rimangono invariati) e dunque $d(P)$ risulta invariante per equiscindibilità, visto che nel processo di suddivisione in pezzi, rotazione dei pezzi e ricomposizione del poliedro, tale quantità non cambia.

Per concludere la dimostrazione del Teorema di Dehn è quindi sufficiente trovare due poliedri dello stesso volume su cui l'invariante assume valori diversi. La scelta più semplice che possiamo fare è quella di prendere un parallelepipedo retto Q e un tetraedro regolare P . Sul parallelepipedo si avrà certamente $d(Q) = 0$ in quanto l'ampiezza degli angoli su tutti gli spigoli è pari a $\pi/2$ e abbiamo imposto $\phi(\pi/2) = 0$. Anche le ampiezze degli angoli sugli spigoli del tetraedro sono tutte uguali tra loro e pari a $\theta = \arccos(1/3)$. Non lo facciamo qui, ma si può dimostrare (con un poco di fatica) che tale angolo θ è incommensurabile con π (cioè $\theta/\pi \notin \mathbb{Q}$) e quindi è possibile scegliere la funzione additiva ϕ in modo tale che $\phi(\theta) \neq 0$. In tal caso si avrà dunque $d(P) \neq 0$ e quindi $d(P) \neq d(Q)$.

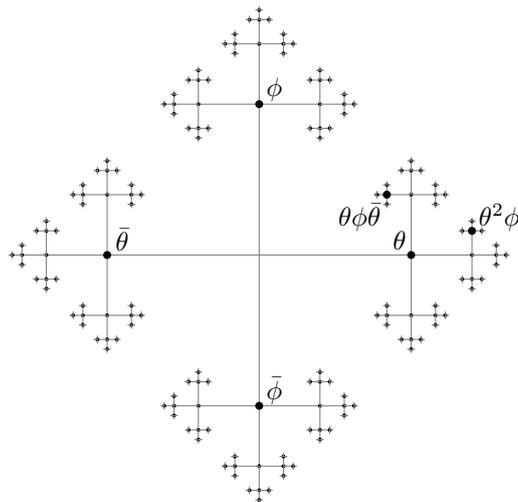


Figura 5: grafo di Cayley del gruppo libero generato da due elementi.

10 Il teorema di Banach-Tarski

Completiamo il nostro percorso con una descrizione della dimostrazione del teorema di Banach-Tarski. Il primo enunciato che dimostreremo è il seguente: la palla unitaria di \mathbb{R}^3 è equidecomponibile a due copie disgiunte di se stessa. Questo sarà il risultato più complesso da dimostrare. Poi con qualche ulteriore facile passaggio riusciremo a dimostrare il risultato più generale: se A e B sono sottoinsiemi di \mathbb{R}^n limitati e con parte interna non vuota e se $n \geq 3$ allora A e B sono equidecomponibili: $A \sim B$.

Questo risultato è evidentemente paradossale, nella misura in cui riteniamo che il volume di un oggetto (supponiamo omogeneo) sia proporzionale alla sua massa e che la massa debba conservarsi in una qualunque suddivisione geometrica dell'oggetto stesso. In effetti questo teorema scaturisce da una ricerca fatta da Banach e Tarski nel tentativo di portare alle estreme conseguenze l'assioma della scelta. Essi volevano infatti verificare se tale assioma fosse in contraddizione con gli altri assiomi di Zermelo-Fraenkel della teoria degli insiemi e riprendendo un risultato di Hausdorff (che vedremo tra poco) sulla decomposizione *paradossale* di una superficie sferica hanno tentato di spingerlo alle estreme conseguenze per cercare di arrivare ad un assurdo.

Quello che hanno trovato non è risultato in contraddizione con gli altri assiomi della teoria degli insiemi ed è, appunto, il Teorema di Banach-Tarski (a volte citato come Teorema di Hausdorff-Banach-Tarski per evidenziare l'importante contributo dato anche da Hausdorff a questo risultato).

La prima parte della dimostrazione del paradosso di Banach-Tarski, come abbiamo già detto, è in realtà dovuta ad Hausdorff. La costruzione di Hausdorff può essere immaginata come una generalizzazione a due dimensioni dell'esempio di Vitali. Si tratta di decomporre la superficie sferica \mathbb{S}^2 (in realtà: un suo sottoinsieme, ci saranno dei *trucioli* da riassorbire) in modo paradossale, ovvero con un numero finito di copie di se stessa.

L'idea è di considerare due rotazioni θ e ϕ . Le rotazioni sono isometrie che mandano la sfera unitaria in se stessa. Richiediamo che queste rotazioni siano *indipendenti* nel senso che se applichiamo in successione le rotazioni θ e ϕ o le loro inverse $\bar{\theta} = \theta^{-1}$ e $\bar{\phi} = \phi^{-1}$ senza mai applicare immediatamente di seguito una rotazione e la sua propria inversa, allora non otteniamo mai l'identità. Detto in altri

termini, il gruppo G delle rotazioni generate da θ e ϕ è *liberamente generato* e può essere rappresentato tramite il grafo frattale che abbiamo già visto in figura 5. Ogni elemento del gruppo G è una rotazione che può essere rappresentata tramite una parola composta da una sequenza della quattro lettere $\theta, \phi, \bar{\theta}, \bar{\phi}$ con la condizione che tale parola non contenga mai consecutivamente una lettera e la sua inversa. L'operazione interna del gruppo corrisponde alla concatenazione delle parole, con l'eventuale cancellazione delle sequenze formate da una lettera e la sua inversa.

Ad esempio due elementi del gruppo G sono $g = \theta\phi\bar{\theta}$ e $h = \theta^2\phi = \theta\theta\phi$. L'elemento gh si ottiene componendo di seguito prima g e poi h , si ottiene quindi $gh = \theta\phi\bar{\theta}\theta\theta\phi = \theta\phi\theta\phi$. Nel grafo rappresentato in figura 5 l'elemento g corrisponde al vertice che si raggiunge partendo dal centro e muovendosi verso est, poi nord e poi ad ovest. Per arrivare al vertice gh si continua ritornando verso est poi ancora verso est e poi a nord.

Una possibile scelta di rotazioni indipendenti è data da due rotazioni dello stesso angolo $\alpha = \arccos \frac{1}{3}$ (lo stesso che abbiamo utilizzato nel teorema di Dehn!) ma con assi di rotazione perpendicolari. Il fatto che il gruppo G generato da queste due particolari rotazioni sia liberamente generato richiede una dimostrazione piuttosto noiosa che tralasciamo (ma si veda [8, Theorem 2.1]).

Mettiamo ora in evidenza, dal punto di vista algebrico, la struttura frattale di questo gruppo G . La radice è l'identità e (sarebbe la parola *vuota*). Gli altri elementi del gruppo possono essere suddivisi in quattro rami, in base a qual è la prima lettera della parola corrispondente: $G_\theta, G_\phi, G_{\bar{\theta}}, G_{\bar{\phi}}$. Dunque abbiamo:

$$G = \{e\} \cup G_\theta \cup G_\phi \cup G_{\bar{\theta}} \cup G_{\bar{\phi}}.$$

Per esprimere la struttura autosimilare del gruppo possiamo osservare che ogni ramo è isomorfo, all'unione degli altri tre. Ad esempio se prendiamo il ramo $G_{\bar{\theta}}$, formato da tutte le parole che iniziano per $\bar{\theta}$ (il ramo di sinistra nella figura) e applichiamo la rotazione θ (cioè aggiungiamo la lettera θ a sinistra della parola e la cancelliamo con la prima $\bar{\theta}$ della parola) otteniamo tutte le parole di G tranne quelle che iniziano per θ (in quanto se la prima lettera era $\bar{\theta}$ la seconda non poteva essere θ). Algebricamente:

$$\theta G_{\bar{\theta}} = \{e\} \cup G_{\bar{\theta}} \cup G_\phi \cup G_{\bar{\phi}}$$

da cui le due decomposizioni (la seconda si ottiene in modo analogo alla prima):

$$G = \theta G_{\bar{\theta}} \cup G_\theta, \quad G = \phi G_{\bar{\phi}} \cup G_\phi.$$

Questa decomposizione di G viene detta *paradossale* in quanto G si decompone in quattro rami che possono essere ricomposti a due a due in modo da ottenere due copie identiche di G (e ci avanza la radice $\{e\}$).

Questo è il nocciolo della dimostrazione del Teorema di Banach-Tarski. Una volta determinata una decomposizione paradossale del gruppo G , sfrutteremo il fatto che questo gruppo è rappresentato fedelmente nelle orbite di quasi tutti i punti della sfera. Poi accorperemo molte orbite assieme, in modo da ricoprire tutta la sfera (o quasi tutta). Infine, visto che gli elementi di G sono isometrie della sfera, noteremo che la decomposizione paradossale di G si traduce in una decomposizione di quasi tutta la sfera. Considerando i coni con vertice nel centro della sfera, otteniamo infine una decomposizione paradossale di quasi tutta la palla unitaria. Si tratterà, infine, di gestire l'insieme numerabile di trucioli che sono rimasti fuori da questa decomposizione e questo sarà un dettaglio tecnico da risolvere tramite assorbimento. Continuiamo quindi ora con la dimostrazione, tenendo in mente la traccia appena delineata.

La decomposizione paradossale del gruppo libero generato da due rotazioni ci permette di ottenere, passando alle orbite dei punti sulla sfera, un insieme frattale

denso in \mathbb{S}^2 . Infatti se prendiamo un punto e guardiamo l'insieme formato dall'immagine di x tramite ognuno degli elementi g del gruppo G otteniamo un insieme che rappresenta geometricamente il gruppo G . I vertici del grafo rappresentato in figura 5 si avvolgono intorno alla sfera.

L'unica accortezza è di scegliere il punto $x \in \mathbb{S}^2$ in modo che non stia su nessuno degli assi delle rotazioni di G (altrimenti alcune rotazioni lascerebbero fisso il punto, e non sarebbero quindi rappresentate fedelmente). Sia dunque Σ' l'insieme formato da tutti i punti fissi degli elementi di G . Essendo G un insieme numerabile anche Σ' è numerabile. Consideriamo inoltre l'insieme $\Sigma = G\Sigma'$ formato da tutti i punti che si possono ottenere da Σ' applicando qualunque rotazione di G . Anche questo insieme è numerabile, perché prodotto di insiemi numerabili. Abbiamo ora un insieme Σ che contiene i punti fissi degli elementi di G e che è invariante rispetto agli elementi di G . Dunque il suo complementare $\Omega = \mathbb{S}^2 \setminus \Sigma$ è un insieme che *non* contiene i punti fissi di G e anch'esso è invariante³.

Preso dunque un qualunque punto $x \in \Omega$, l'insieme $G(x) = \{g(x) : g \in G\}$ (chiamato *orbita* di x) è effettivamente in corrispondenza biunivoca con G . Infatti se ci fossero $g, h \in G$ tali che $g(x) = h(x)$ allora x sarebbe un punto fisso di $g^{-1}h$ cosa che abbiamo esplicitamente escluso. L'intera sfera bucherellata Ω può essere suddivisa in una partizione di orbite di punti (due punti x e y stanno sulla stessa orbita se $G(x) = G(y)$ cioè se esiste una rotazione $g \in G$ tale che $y = g(x)$). Tramite l'assioma della scelta, possiamo determinare un insieme Δ di rappresentanti di queste orbite, cioè Δ conterrà uno e un solo punto di ogni orbita. Dunque avremo una partizione di Ω in orbite di punti di Δ :

$$\Omega = \bigcup_{x \in \Delta} G(x) = \{g(x) : x \in \Delta, g \in G\}$$

ovvero Ω stesso può essere visto come una rappresentazione geometrica del gruppo G :

$$\Omega = \bigcup_{g \in G} g(\Delta)$$

e potrà quindi essere decomposto in quattro rami, seguendo la decomposizione di G :

$$\Omega = \Delta \cup G_\theta(\Delta) \cup G_\phi(\Delta) \cup G_{\bar{\theta}}(\Delta) \cup G_{\bar{\phi}}(\Delta) \quad (8)$$

e la decomposizione *paradossale* di G diventa una decomposizione *paradossale* di Ω :

$$\Omega = \theta G_{\bar{\theta}}(\Delta) \cup G_\theta(\Delta), \quad \Omega = \phi G_{\bar{\phi}}(\Delta) \cup G_\phi(\Delta).$$

Visto che θ e ϕ sono isometrie abbiamo qui ottenuto una decomposizione di Ω in cinque pezzi (8) che ricomposti a due a due tramite rotazioni, ci permettono di ottenere due copie identiche di Ω con l'avanzo del pezzo Δ . Ovviamente si intende che una delle due copie di Ω può essere resa disgiunta dall'altra mediante una opportuna traslazione, e lo stesso vale per l'insieme Δ che può essere traslato in modo da non sovrapporsi alle due copie di Ω .

E' chiaro ora che questa decomposizione (di quasi tutta la sfera) può essere immediatamente estesa a quasi tutta la palla. Basta pensare che la palla piena è l'unione dei suoi raggi aperti⁴, ognuno dei quali identificato da un punto sulla superficie sferica. Visto che le due rotazioni θ e ϕ hanno entrambe centro nel centro

³Se un punto x di Ω andasse in Σ tramite una rotazione $g \in G$, allora esisterebbe un'altra rotazione $h \in G$ per cui $hg(x) \in \Sigma'$. Ma allora x sarebbe raggiungibile da Σ' tramite la rotazione inversa $g^{-1}h^{-1}$ e dunque dovrebbe stare in Σ .

⁴Dobbiamo escudere il centro dai raggi, per fare sì che i raggi siano disgiunti. Possiamo tenere o meno l'altro estremo del raggio a seconda che vogliamo decomporre una palla aperta oppure chiusa.

della palla, si vede chiaramente che la decomposizione paradossale di Ω può essere estesa ad una decomposizione paradossale di quasi tutta la palla unitaria \mathbb{B}^3 . Se chiamiamo C il cono costruito su Ω (senza il vertice) e se chiamiamo D il cono costruito su Δ , si ottiene infatti:

$$C = D \cup G_\theta(D) \cup G_\phi(D) \cup G_{\bar{\theta}}(D) \cup G_{\bar{\phi}}(D)$$

ma anche

$$C = \theta G_{\bar{\theta}}(D) \cup G_\theta(D), \quad C = \phi G_{\bar{\phi}}(D) \cup G_\phi(D).$$

Abbiamo quindi ottenuto una decomposizione di un insieme C corrispondente a quasi tutta la palla \mathbb{B}^3 in cinque pezzi, quattro dei quali possono essere ricomposti (mediante isometrie) in modo da riottenere due copie identiche di tutto C e un pezzo D avanza. Chiamando C' una copia traslata di C abbiamo quindi:

$$C \cup C' \sim C \setminus D \subseteq C \subseteq C \cup C'$$

che significa che $C \cup C' \preceq C$ e $C \preceq C \cup C'$. Dunque applicando il teorema di Cantor-Bernstein nella variante relativa alle equidecomposizioni, possiamo concludere che

$$C \cup C' \sim C.$$

Siamo arrivati ad un punto in cui l'aspetto paradossale del risultato è già evidente. Salvo un insieme numerabile di raggi, l'intera palla unitaria può essere spezzata e ricomposta per ottenere due copie di sé.

Ora ci dedichiamo al problema di *riassorbire* quell'insieme numerabile di raggi $\Gamma = \mathbb{B}^3 \setminus C$ che sono rimasti esclusi dalla decomposizione (se ci ricordiamo, questi erano gli assi delle due rotazioni θ e ϕ e tutte le loro immagini tramite il gruppo G delle rotazioni generate da θ e ϕ).

L'idea è ora quella di sfruttare ancora una volta il paradosso dell'infinito (l'hotel Hilbert). Vogliamo mostrare che esiste una rotazione τ tale che $\tau^n(\Gamma)$ sia disgiunto da $\tau^m(\Gamma)$ per ogni $n \neq m$. Per determinare un tale τ possiamo innanzitutto fissare a piacere un asse di rotazione. Poi consideriamo tutte le coppie di raggi di Γ e se c è una rotazione intorno all'asse fissato che manda un raggio della coppia nell'altro, consideriamo l'angolo di tale rotazione. L'insieme di tutti questi angoli è numerabile (perché tale è Γ e quindi anche l'insieme di coppie di Γ) e dunque esisterà un angolo incommensurabile con tutti questi. Prendiamo dunque una rotazione τ intorno all'asse che abbiamo scelto dell'angolo che abbiamo scelto. Questa rotazione τ avrà la proprietà richiesta, che si traduce nel fatto che l'orbita tramite τ dell'insieme Γ , è una unione disgiunta:

$$X = \bigcup_{n \in \mathbb{N}} \tau^n(\Gamma).$$

Ora X può essere decomposto in due pezzi disgiunti, di cui il secondo è isometrico a X stesso:

$$X = \Gamma \cup \bigcup_{n \in \mathbb{N}} \tau^{n+1}(\Gamma) = \Gamma \cup \tau(X).$$

Abbiamo quindi mostrato che X è equidecomponibile all'unione di Γ con una copia di X stesso. Questo significa che $\mathbb{B}^3 \setminus \{0\} = C \cup \Gamma$ è equidecomponibile a C , in quanto

$$C \cup \Gamma = (C \setminus \tau(X)) \cup \tau(X) \cup \Gamma = (C \setminus \tau(X)) \cup X \sim (C \setminus \tau(X)) \cup \tau(X) = C.$$

In questo modo abbiamo finalmente ottenuto una decomposizione paradossale di $\mathbb{B}^3 \setminus \{0\}$ in due copie identiche di se stesso. Ci rimane fuori solamente il centro della sfera, che per essere duplicato necessita di un procedimento di riassorbimento

simile a quello utilizzato per l'insieme Γ . Si tratta questa volta di considerare una rotazione con un asse che, per la prima volta, non passa dall'origine, con un angolo di rotazione irrazionale rispetto all'angolo giro, e tale che l'orbita dell'origine stia tutta all'interno della palla unitaria (cioè l'asse deve distare dall'origine per meno di mezza unità). L'orbita dell'origine sarà quindi un insieme numerabile di punti che, utilizzando la rotazione scelta, sarà equidecomponibile ad una copia di se stesso più l'origine. Mettendo in gioco anche questo nuovo *pezzo* della decomposizione abbiamo finalmente ottenuto una decomposizione di tutta la palla unitaria \mathbb{B}^3 in un numero finito di pezzi che ricomposti tramite isometrie danno origine a due copie identiche della palla iniziale.

11 Banach-Tarski: generalizzazioni

Il teorema di Banach-Tarski nella sua formulazione standard lo abbiamo dimostrato: la palla unitaria è equidecomponibile a due copie di se stessa. Questo risultato è già sufficientemente paradossale e non servirebbe andare a cercare ulteriori generalizzazioni. D'altra parte, avendone visto la dimostrazione, non ci costa molta fatica enunciare una forma apparentemente molto più generale: tutti i sottoinsiemi dello spazio che siano limitati e con parte interna non vuota, sono tra loro equidecomponibili.

Questa generalizzazione si ottiene abbastanza facilmente partendo dal risultato ottenuto per le palle, e utilizzando il teorema di Cantor-Bernstein nella versione adattata alla equidecomposizione.

Infatti sappiamo che una singola palla è equidecomponibile a due palle disgiunte. Ripetendo il procedimento su una delle due palle, sappiamo che essa stessa è equidecomponibile ad altre due palle disgiunte. Iterando il procedimento e applicando il principio di induzione, arriviamo a dire che una singola palla è equidecomponibile ad un insieme formato da n palle disgiunte, per qualunque intero positivo n . Inoltre il raggio di queste palle non è rilevante, fintantoché tutte le palle considerate hanno lo stesso raggio.

Ora se A è un qualunque insieme limitato e $\rho > 0$ è un numero fissato, anche molto piccolo, sappiamo che A può essere ricoperto da un numero finito di palle di raggio ρ (non necessariamente disgiunte tra loro). Dunque è chiaro che A è equidecomponibile ad un sottoinsieme di una unione finita di palle di raggio ρ . Ma l'unione finita di palle di raggio ρ sono equidecomponibili ad una singola palla di raggio ρ , per quanto visto prima. Dunque A risulta essere equidecomponibile ad un sottoinsieme di una palla di raggio ρ . Se B è un insieme con parte interna non vuota esisterà, per definizione, un $\rho > 0$ tale che la palla di raggio ρ , opportunamente traslata, sia un sottoinsieme di B . Risulta quindi che se A è limitato e B ha parte interna non vuota, allora A è equidecomponibile ad un sottoinsieme di B : $A \preceq B$.

Invertendo i ruoli di A e B , se entrambi sono limitati e con parte interna non vuota, abbiamo quindi dimostrato che ognuno dei due è equidecomponibile ad un sottoinsieme dell'altro. Applicando allora la versione del teorema di Cantor-Bernstein adattata all'equidecomponibilità, si ottiene che i due insiemi A e B sono tra loro equidecomponibili, come volevamo dimostrare.

Può essere utile anche osservare che la dimensione dello spazio ambiente può essere aumentata senza inficiare la dimostrazione del teorema di Banach-Tarski. Infatti le rotazioni indipendenti θ e ϕ possono essere estese a qualunque spazio euclideo di dimensione n se $n \geq 3$. Non si trovano invece in \mathbb{R}^2 in quanto le rotazioni concentriche del piano formano un gruppo commutativo all'interno del quale non è possibile trovare un sottogruppo omeomorfo al prodotto libero $\mathbb{Z} * \mathbb{Z}$ in quanto la condizione $\theta\phi = \phi\theta$ non è prescindibile. In effetti, come già accennato, è possibile definire una funzione di insieme additiva (ma non σ -additiva) definita su

tutti i sottoinsiemi limitati di \mathbb{R}^2 a valori reali e non identicamente nulla. Questo ci dice che non è possibile trovare una decomposizione paradossale in questo ambiente.

Riferimenti bibliografici

- [1] V. G. Boltianskii. *Hilbert's third problem*. Winston, 1978.
- [2] K. J. Falconer. *The geometry of fractal sets*. Cambridge University Press, 1985.
- [3] G. Frederickson. *Dissections: Plane & Fancy*. Cambridge University Press, 1997.
- [4] G. H. Moore. *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*. Studies in the History of Mathematics and Physical Sciences 8. Springer-Verlag New York, 1982.
- [5] K. Shiga, T. Sunada. *A Mathematical Gift, III - the interplay between topology, functions, geometry, and algebra*. AMS, 2005.
- [6] R. M. Solovay. *A model of set-theory in which every set of reals is Lebesgue measurable*. *Annals of Mathematics* **92**, 1—56, 1970.
- [7] T. Tao. *The Banach-Tarski paradox*. <https://www.math.ucla.edu/~tao/preprints/Expository/banach-tarski.pdf>.
- [8] S. Wagon. *The Banach-Tarski paradox*. Cambridge University Press, 1985.