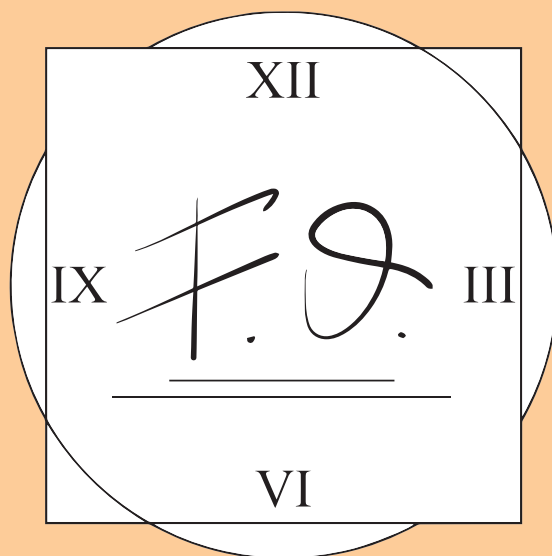


---

F. S. ZARNO

---

Resoconti di una giornata di seminari  
4 Maggio 2022

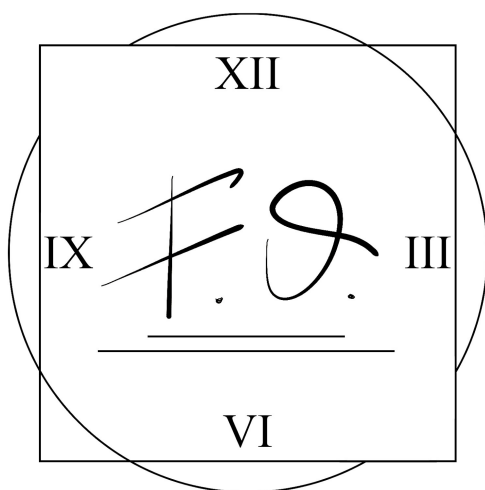




# FUORI ORARIO

Resoconti di una giornata di seminari

4 Maggio 2022



Fuori Orario  
Resoconti di una giornata di seminari

4 Maggio 2022  
Dipartimento di Matematica  
Università degli Studi di Milano

Versione definitiva: settembre 2022

Impaginazione in  $\text{\LaTeX}$  e progetto grafico a cura di Francesco Pagliarin

## INDICE

INTRODUZIONE	1
A PROPOSITO DELLA $\phi$ DI EULERO <i>Marco Della Penna</i>	3
COSA POSSIAMO DIRE DI SPAZI METRICI NUMERABILI PRIVI DI PUNTI ISOLATI? <i>Matteo Anoffo</i>	7
I NUMERI NASCOSTI NELLA SCALA MUSICALE <i>Elisa Pedrini</i>	9
MUSICA E MATEMATICA: UN LEGAME INDISSOLUBILE <i>Edward Kevin Arana Medina</i>	21
CHE CHAOS! <i>Urmila Bosisio</i>	27
COSA C'ENTRA IL CALCOLO DELLE VARIAZIONI CON L'OTTICA GEOMETRICA? <i>Alessia Severgnini</i>	33
STRANEZZE $p$ -ADICHE <i>Simone Puleio</i>	41
UNA PUNTA DI IPOTESI DEL CONTINUO <i>Paolo Boldrini</i>	47
LA LEGGE DI BENFORD – OVVERO: COME HO IMPARATO AD EVITARE FRODI CON LA MATEMATICA <i>Filippo Beretta</i>	55
DA DIDONE A GROMOV: IL PROBLEMA ISOPERIMETRICO E LA SUA SOLUZIONE <i>Beatrice Ostorero Vinci</i>	61
DIMOSTRAZIONI PROBABILISTICHE DI RISULTATI ANALITICI <i>Gabriele Cassese</i>	69



## INTRODUZIONE

Caro lettore o cara lettrice,

Sembra passato pochissimo tempo da quando l'anno scorso mi sono ritrovata a scrivere l'introduzione per la raccolta di seminari della quinta edizione, eppure è già passato un anno, quasi senza che me ne accorgessi.

Ho preso decisioni su quella che sarà la mia vita per i prossimi due anni, e un po' a malincuore, queste mi portano lontano da Via Saldini e da tutto quello che la riguarda. Pertanto, queste righe saranno una sorta di saluto a Fuori Orario, a chi lo ha reso e lo rende possibile ogni anno e al Dipartimento che mi ha fatto conoscere tutto questo.

Fuori Orario ha arricchito i miei tre anni passati in università: mi ha fatto crescere, dal punto di vista matematico e non, mi ha messo in situazioni nuove, che si sono rivelate sfide stimolanti e occasioni di incontro e di confronto con molte persone. Ovviamente in queste situazioni è necessario giocare attivamente la propria parte, ma sono abbastanza sicura che un contesto e un'iniziativa diversi non avrebbero portato allo stesso risultato. Ho conosciuto tanta gente grazie a Fuori Orario, ragazzi e ragazze che con me hanno preso a cuore l'iniziativa e hanno deciso di partecipare, dando il loro contributo. Vorrei ringraziare chi ha condiviso con me questa avventura, per uno, due o per tre anni: dalle riunioni per organizzare alla gestione delle giornate di seminari, c'è stato un buon lavoro di squadra, senza il quale probabilmente sarebbe stato tutto molto più difficile.

Quest'anno in particolare, abbiamo affrontato una sfida non indifferente: organizzare la sesta edizione in presenza dopo due anni di seguito di edizioni online. C'erano tante incognite, molti più aspetti da prendere in considerazione, e tutti noi avevamo fatto esperienza solo della componente telematica dell'iniziativa; tuttavia, nonostante qualche difficoltà, il team di organizzatori è riuscito a riportare Fuori Orario nelle aule, recuperando la dimensione originale della giornata di seminari. Anche gli speakers e il pubblico hanno accolto con entusiasmo la possibilità di ritrovarsi fisicamente a presentare e ascoltare curiosità sulla matematica: nell'arco della giornata si sono svolti undici interventi a cui hanno partecipato circa un centinaio di studenti e studentesse, spinti dalla curiosità e dalla felicità di un ritorno alla normalità tanto atteso.

Sono contenta che questa sia l'edizione con cui saluto Fuori Orario. Da un lato mi ha permesso di collaborare con le persone dal vivo, di provare a parlare ad un pubblico fisico, e non davanti ad uno schermo, tra le pareti della mia stanza; dall'altro

quest'anno più degli altri ho avuto la possibilità di vedere mettersi in gioco persone nuove, tra cui anche persone a me vicine, e trovo che l'essere riusciti a coinvolgerle, trasmettendo e condividendo con loro quello che è lo spirito di questa iniziativa, sia bello.

A questo punto, credo di potermi avviare verso la conclusione di questi saluti, per quanto sia sempre difficile. Mi spiace un po' non poter continuare a partecipare tanto attivamente a quella che mi auguro sarà la continua crescita ed evoluzione di Fuori Orario, ma ho visto tante nuove idee cominciare a svilupparsi e tanta gente con entusiasmo e voglia di fare, e questo mi fa pensare di lasciare tutto in buone mani. Spero che non ci siano più pandemie a portare via quello che in questa sesta edizione abbiamo riavuto, ma soprattutto, spero che nelle prossime edizioni ci sia sempre più varietà di argomenti e livelli, vengano coinvolti sempre più i piccoli del primo anno e si riesca ad avere più parità tra ragazzi e ragazze tra gli speakers, perché, anche se da fuori può sembrare tale, Fuori Orario non è una vetrina o un piedistallo per chi sa già tutto delle materie più difficili, ma un'occasione per divertirsi e condividere una matematica diversa.

Mi auguro davvero che chi si ritroverà tra le mani Fuori Orario, continui ad impegnarsi con strategie e cambiamenti perché passi sempre più questa idea dell'iniziativa e che tante persone si sentano accolte e spronate, come è successo a me.

Lasciando Via Saldini, e quindi Fuori Orario, mi allontano da un ambiente in cui mi sento sicura, che mi ha permesso di essere me stessa, di fare incontri speciali e che ha contribuito a formare la persona e la matematica che sto diventando. Cambiare università e paese vuol dire andare verso qualcosa che non si conosce, consapevoli di quello a cui si sta rinunciando, e, specialmente nel momento in cui la partenza si avvicina e ci si ferma a pensare, si sente un leggero nodo alla gola. La verità però è che o adesso o tra un paio d'anni, avrei dovuto salutare tutto questo in ogni caso e, anche se organizzare Fuori Orario non sarà più mio compito e i corridoi del Dipartimento non saranno più la normalità, potrò comunque continuare a seguire le giornate di seminari, e tutto ciò che è stato in questi tre anni verrà con me, come ricordo, come sicurezza acquisita e molto altro.

Ora è davvero arrivato il momento di salutarci, caro lettore o cara lettrice, e di lasciarti libero o libera di sfogliare queste pagine e di soffermarti su ciò che più ti incuriosisce.

Con un po' di malinconia, ti auguro buona lettura.

Beatrice Ostorero Vinci



# A PROPOSITO DELLA $\phi$ DI EULERO

MARCO DELLA PENNA

## 1. Introduzione

L'obiettivo dell'intervento è di sfruttare una disuguaglianza assai complessa ma altrettanto affascinante riguardante la  $\phi$  di Eulero per scavare nel mondo dell'algebra dei numeri interi e, soprattutto, dei numeri primi. Il testo del problema 6 della sezione "Number Theory" delle "IMO Shortlisted" recita così: "Esiste una costante  $C$  tale che vale

$$\frac{\phi(d(n))}{d(\phi(n))} \leq C$$

$\forall n \geq 1$ ?"

In particolare  $d(n)$  indica il numero di divisori di  $n$  ed è ben noto che se  $n = \prod_{i=1}^k p_i^{\alpha_i}$  (una facile induzione su  $k$  lo dimostra),  $d(n) = \prod_{i=1}^k (\alpha_i + 1)$ .  $\phi(n)$  invece indica il numero di interi positivi coprimi con  $n$  e minori di  $n$  ed è ben noto che  $\phi(n) = \prod_{i=1}^k (p_i^{\alpha_i} - p_i^{\alpha_i-1}) = \prod_{i=1}^k p_i^{\alpha_i-1} (p_i - 1)$ .

## 2. Come approcciare il problema

Il modo giusto, se ne esiste uno, per affrontare un problema di questo tipo è "iniziare a provare": è importante farsi un'idea!

Per esempio vediamo che se  $n$  fosse un numero primo si avrebbe  $\phi(d(n)) = \phi(2) = 1 \leq d(n-1) = d(\phi(n))$  ma, già più in generale, se  $n$  fosse la  $h$ -esima potenza di un primo  $p$ ,  $\phi(d(n)) = \phi(h+1) \leq 2h \leq 2d(p^{h-1}(p-1)) = 2d(\phi(n))$ . Proviamo ora ad affacciarci al caso generale:

$$\phi(d(n)) = \phi\left(\prod_{i=1}^k (\alpha_i + 1)\right) \tag{1}$$

$$d(\phi(n)) = d\left(\prod_{i=1}^k p_i^{\alpha_i-1} (p_i - 1)\right) \tag{2}$$

Risulta dunque evidente, guardando solo il contributo dei  $p_i^{\alpha_i-1}$ , che  $d(\phi(n)) \geq \prod_{i=1}^k \alpha_i$  e tale produttoria assomiglia molto all'argomento della  $\phi$  dell'equazione 1. Contiamo, dunque, di riuscire a minorare ulteriormente  $d(\phi(n))$  attraverso il contributo dei  $p_i - 1$  e di poter invece maggiorare  $\phi(d(n))$  attraverso l'azione di  $\phi$ , appunto, che riduce il

proprio argomento. Tutto sembrerebbe portare a pensare che esista una tale  $C$  ma vediamo ora che la differenza apparentemente piccola tra gli  $\alpha_i$  e i corrispondenti  $\alpha_i + 1$  potrebbe essere molto più difficile da colmare del previsto, dopotutto è sempre un IMO 6! :)

Partiamo provando a maggiorare la (1)...purtroppo dobbiamo fermarci molto presto: all'apparenza la  $\phi$  dell'equazione (1) sembra molto difficile da stimare ma una scelta abbastanza intuitiva lo farà diventare banale: poniamo  $\alpha_1 = \dots = \alpha_k = q - 1$  con  $q$  primo. Risulta, dunque, ovvio che  $\phi(d(n)) = \phi(q^k) = q^{k-1}(q - 1)$ : in poche parole non abbiamo ottenuto niente, in quanto siamo riusciti a diminuire soltanto uno dei  $k$   $\alpha_i + 1$ !!! Notiamo, inoltre, che la nostra scelta arbitraria non influisce più di tanto sull'equazione (2) poiché, essendo gli  $\alpha_i - 1$  all'esponente, la loro fattorizzazione in primi non è particolarmente interessante ma conta soltanto quanto sono grandi: questa è una grande svolta nel problema, non siamo più così sicuri che esista una tale  $C$ .

Concentriamoci ora sulla (2): ci piacerebbe essere in grado di dire qualcosa in più a riguardo del contributo dei vari  $p_i - 1$  ma per fare ciò è fondamentale la fattorizzazione in primi di essi. In particolare vorremmo sapere se ci sono (e in che quantità) dei primi diversi dai vari  $p_i$ . Per fare ciò basta scegliere di prendere tutti i primi  $p_i$  fino a un certo numero  $2N$ : in questo modo sapremo che ogni  $p_i - 1$  ha come fattori soltanto i  $p_i$  poiché è minore di  $2N$ . In realtà possiamo dire anche qualcosa in più: tralasciando  $p = 2$  (che non scambussola più di tanto le carte in tavola) ogni  $p_i - 1$  è pari e, in particolare, avrà fattori minori di  $N$ . Possiamo, dunque, sapere con certezza i coefficienti dei primi compresi tra  $N$  e  $2N$ : vediamo che impatto hanno questi ultimi sul risultato. Iniziamo col caso più semplice: scegliamo tutti gli esponenti di questi  $s$  primi uguali a 1, allora il contributo che essi danno alla (1) è un fattore moltiplicativo pari a  $\phi(2^s) = 2^{s-1}$ . Nella (2), invece, come già detto, il loro contributo andrà ad aumentare i coefficienti dei primi minori di  $N$ . Ecco, dunque, un altro passo decisivo verso l'idea corretta, ovvero che in effetti non può esistere una tale costante  $C$ .

Direi che come approccio al problema ci siamo spinti fin troppo oltre, passiamo alla soluzione ufficiale.

### 3. Il comportamento dei reciproci dei primi

Prima di tutto è opportuno mostrare che  $s$  può essere reso arbitrariamente grande i.e. che, a patto di scegliere  $N$  sufficientemente grande, posso trovare un numero sempre più grande di primi compresi fra  $N$  e  $2N$ . Questa è una diretta conseguenza della divergenza della serie dei reciproci dei primi, che, per completezza, dimostriamo.

LEMMA.

$$\sum_{p \in P} \frac{1}{p} = \infty$$

dove  $P$  indica l'insieme dei numeri primi.

*Dimostrazione.* Basta osservare che vale la seguente catena di relazioni:

$$\begin{aligned}
\sum_{p \in P} \frac{1}{p} &= \sum_{p \in P} \frac{1}{p-1} + \sum_{p \in P} \frac{1}{p} - \sum_{p \in P} \frac{1}{p-1} = \sum_{p \in P} \frac{1}{p-1} - \sum_{p \in P} \frac{1}{p(p-1)} \geq \\
&\geq \sum_{p \in P} \frac{1}{p-1} - \sum_{n=2}^{\infty} \frac{1}{n(n-1)} = \sum_{p \in P} \frac{1}{p-1} + \sum_{n=2}^{\infty} \frac{1}{n} - \sum_{n=2}^{\infty} \frac{1}{n-1} = \sum_{p \in P} \frac{1}{p-1} - 1 \geq \\
&\geq \sum_{p \in P} \log\left(1 + \frac{1}{p-1}\right) - 1 = \log\left(\prod_{p \in P} 1 + \frac{1}{p-1}\right) - 1 = \log\left(\prod_{p \in P} \frac{1}{1 - \frac{1}{p}}\right) - 1 = \\
&= \log\left(\prod_{p \in P} \sum_{k=0}^{\infty} \frac{1}{p^k}\right) - 1 = \log\left(\sum_{n=1}^{\infty} \frac{1}{n}\right) - 1 = \infty
\end{aligned}$$

□

Adesso avete tutti gli strumenti a disposizione per dimostrare che  $s$  può tendere all'infinito e, perché no, sfruttando i suggerimenti della sezione 2 potete anche risolvere il problema iniziale. Vi invito, dunque, a provare a lungo ad arrivare alla soluzione e solo dopo numerosi tentativi vi concedo di arrendervi e di guardare la soluzione ufficiale del problema che è riportata di seguito.

#### 4. Dimostrazione

Fissiamo  $N > 1$ , siano  $p_1, \dots, p_k$  tutti i primi tra 1 e  $N$  e  $p_{k+1}, \dots, p_{k+s}$  tutti i primi tra  $N+1$  e  $2N$ . Siccome per  $j \leq k+s$  tutti i divisori primi di  $p_j - 1$  non eccedono  $N$ , abbiamo

$$\prod_{j=1}^{k+s} (p_j - 1) = \prod_{i=1}^k p_i^{c_i}$$

per alcuni esponenti fissati  $c_1, \dots, c_k$ . Scegliamo un numero primo enorme  $q$  e consideriamo il numero

$$n = (p_1 \cdot \dots \cdot p_k)^{q-1} \cdot (p_{k+1} \cdot \dots \cdot p_{k+s})$$

Dunque

$$\phi(d(n)) = \phi(q^k \cdot 2^s) = q^{k-1} (q-1) 2^{s-1}$$

e

$$d(\phi(n)) = d((p_1 \cdot \dots \cdot p_k)^{q-2} \prod_{i=1}^{k+s} (p_i - 1)) = d\left(\prod_{i=1}^k p_i^{q-2+c_i}\right) = \prod_{i=1}^k (q-1+c_i)$$

quindi

$$\frac{\phi(d(n))}{d(\phi(n))} = \frac{q^{k-1} (q-1) 2^{s-1}}{\prod_{i=1}^k (q-1+c_i)} = 2^{s-1} \cdot \frac{q-1}{q} \cdot \prod_{i=1}^k \frac{q}{q-1+c_i}$$

che può essere reso arbitrariamente vicino a  $2^{s-1}$  scegliendo  $q$  abbastanza grande. Rimane da mostrare che  $s$  può essere arbitrariamente grande i.e. che possono esserci tanti primi a piacere tra  $N$  e  $2N$ .

Questo segue dal lemma precedente infatti, se per qualche costante  $C$  ci fossero al massimo  $C$  primi tra  $2^l$  e  $2^{l+1}$  avremmo che

$$\sum_{p \in P} \frac{1}{p} = \sum_{l=0}^{\infty} \sum_{p \in P \cap [2^l, 2^{l+1})} \frac{1}{p} \leq \sum_{l=0}^{\infty} C = C \cdot \sum_{l=0}^{\infty} 1 = \infty.$$

## 5. Un altro metodo

Nelle soluzioni ufficiali delle IMO (cf. link a fine pagina) viene riportato anche un procedimento molto più complesso ed avanzato ma anche più affascinante in quanto sfrutta il celebre teorema dei numeri primi:

$$\pi(n) \sim \frac{n}{\log n}$$

dove  $\pi(n)$  è la funzione che associa ad  $n$  il numero di primi minori o uguali a  $n$ . Di seguito riportiamo un po' di hint per chi volesse provare a risolvere il problema anche con questo secondo metodo.

1. Consideriamo il numero  $n = \prod_{p \in P \wedge p \leq \pi(m)} p$ : è ovvio che sia  $\phi(d(n))$  ma cosa dire a riguardo dell'altro termine della frazione? Cominciamo stimando  $\phi(n)$  che si scrive di nuovo come prodotti di primi minori di  $\frac{m}{2}$ :  $\phi(n) = \prod_{s=1}^{\pi(\frac{m}{2})} q_s^{\alpha_s}$ . Iniziamo concentrandoci sui primi maggiori di  $\sqrt{m}$ : mostrate che

$$\sum_{s: q_s > \sqrt{m}} (1 + \alpha_s) \leq \pi(m) + \pi\left(\frac{m}{2}\right)$$

e deducetene che

$$\prod_{s: q_s > \sqrt{m}} (1 + \alpha_s) \leq \exp\left(\frac{\pi(m) + \pi(\frac{m}{2})}{e}\right)$$

2. Mostrate ora invece che

$$\prod_{s=1}^{\pi(\sqrt{m})} (1 + \alpha_s) \leq m^{2\sqrt{m}}$$

3. Sfruttando il teorema dei numeri primi mostrate che

$$\limsup_{m \rightarrow \infty} \frac{\log(d(\phi(n)))}{m/\log m} \leq \frac{3}{2e}$$

$$\liminf_{m \rightarrow \infty} \frac{\log(\phi(d(n)))}{m/\log m} = \log 2$$

4. Dedurne la tesi.

## Bibliografia

[I] PAGINE 81-82, <https://www.imo-official.org/problems/IMO2020SL.pdf>

# COSA POSSIAMO DIRE DI SPAZI METRICI NUMERABILI PRIVI DI PUNTI ISOLATI?

MATTEO ANOFFO

## 1. Il teorema di Sierpinsky

Uno spazio metrico è una coppia  $(X, d)$  dove  $X$  è un insieme e  $d$  è una funzione, detta metrica, per la quale valgano le seguenti proprietà:

- $d(x, y) > 0 \iff x \neq y$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, y) \leq d(x, z) + d(y, z)$ , ovvero vale la disuguaglianza triangolare

Nel nostro caso prenderemo in considerazione spazi metrici numerabili, ossia quelli in cui la cardinalità del nostro insieme  $X$  è numerabile, e privi di punti isolati, ovvero spazi metrici nei quali, scelto un qualsiasi punto  $x$  e una qualsiasi distanza  $r$ , esiste sempre un punto  $y$ , diverso da  $x$ , appartenente all'intorno di  $x$  di raggio  $r$ .

Il teorema che andremo ora a enunciare, e quindi dimostrare, legherà ogni spazio metrico di tale tipo.

**TEOREMA 1.** *Due spazi metrici numerabili privi di punti isolati sono omeomorfi tra di loro, ossia esiste una funzione biettiva continua con inversa continua da uno spazio metrico all'altro.*

*Dimostrazione.* Siano  $(X, d_x)$  e  $(Y, d_y)$  due spazi metrici di tale tipo. Siano  $X = \{x_1, x_2, \dots\}$  e  $Y = \{y_1, y_2, \dots\}$  con  $x_m \neq x_n$  se  $m \neq n$ , e analogo per  $Y$ . Poniamo  $D = \{d_x(x_m, x_n) : m \neq n\} \cup \{d_y(y_m, y_n) : m \neq n\}$ , così che  $D$  è un insieme numerabile di numeri positivi. Per ogni sottoinsieme non vuoto  $S \subset X$  definiamo  $next(S) = x_k$ , con  $k = \min\{i : x_i \in S\}$ , e in modo uguale per  $Y$ . Per  $r > 0$  sia

$$B(S, r) = B_r(next(S)) \quad e \quad G(S, r) = S \setminus B(S, r)$$

Introduciamo ora lo "splitter": per ogni sottoinsieme aperto non vuoto  $S \subset X$ , o  $S \subset Y$ ,  $\exists r > 0$  tale che  $B(S, r)$  e  $G(S, r)$  dividono  $S$  in due sottoinsiemi aperti, non vuoti e disgiunti.

Per vedere questo, poniamo  $x = next(S)$ . Sappiamo per ipotesi che  $x$  non è isolato, esiste quindi  $y \in S$  tale che  $y \neq x$ . Scegliamo  $r \in (0, d_x(x, y)) \setminus D$  piccolo abbastanza da essere  $B(S, r) \subset S$ . Allora  $y \in S \setminus B(S, r) = G(S, r)$ , e poiché  $r \notin D$  abbiamo  $G(S, r) = \{u \in S : d_x(x, u) > r\}$ .  $G(S, r)$  è quindi aperto.

Possiamo usare lo splitter in maniera ricorsiva per generare una sequenza di partizioni

di  $X$  e  $Y$  in insiemi aperti.

Sia  $W_G = \{G, GB, GG, GBB, GBG, GGB, GGG, \dots\}$  l'insieme degli aperti che possiamo generare, in modo che ogni parola inizi con  $G$ . Per  $t \in W_G$  definiamo  $|t|$  la lunghezza di  $t$ , ovvero il numero di lettere di  $t$ , e  $t$  è detta *blue* (rispettivamente *green*) se  $t$  finisce con  $B$  (rispettivamente  $G$ ).

Siano  $U_G = X$  e  $V_G = Y$ . Dimostriamo ora che  $U_t$  e  $V_t$  sono definiti per ogni  $t \in W_G$  con  $|t| \geq 1$ .

Procediamo per induzione sulla lunghezza di  $|t|$ . Applichiamo lo splitter a ogni  $U_t$  e a ogni  $V_t$  con  $|t| = n$  (ve ne sono  $2^{n-1}$  per ciascuno). Poniamo quindi  $U_t \wedge_B = B(U_t, r_{n+1})$  e  $U_t \wedge_G = G(U_t, r_{n+1})$ , e analogamente per  $V_t$ , dove  $r_{n+1}$  è lo stesso per ogni insieme all'iterazione  $n + 1$  dello splitter.  $U_t$  e  $V_t$  sono quindi definiti per ogni  $t \in W_G$ .

Siamo ora pronti a definire l'omomorfismo che stavamo cercando.

Sia  $h : X \rightarrow Y$  definita come segue

$$h(\text{next}(U_t)) = \text{next}(V_t), \quad \text{con } t \in W_G$$

Osserviamo quindi che quando  $x \in X$  è tale che  $x = \text{next}(U_t)$ , per una qualche  $t$ ,  $x$  diventa il centro di  $U_s$ , con  $s \in \{t \wedge B, t \wedge BB, \dots\}$ , e questo crea quindi una sequenza decrescente di bolle di centro  $x$  e raggio  $r \rightarrow 0$ .

Questo assicura che ogni  $x$  è associata ad un'unica *green*  $t \in W_G$ , e uguale per ogni  $y \in Y$ . La nostra funzione  $h$  è quindi una biezione tra  $X$  e  $Y$ . La continuità è invece garantita dall'ipotesi  $h(U_t) = V_t$  per ogni  $t$ ; quindi  $h$  soddisfa il criterio  $\varepsilon, \delta$  per le bolle *blue*.  $\square$

L'esempio più semplice di spazio metrico numerabile privo di punti isolati è sicuramente  $\mathbb{Q}$ . Come immediata conseguenza del teorema appena dimostrato abbiamo infatti il seguente corollario.

**COROLLARIO 2.** *Ogni spazio metrico numerabile è omeomorfo a un sottospazio di  $\mathbb{Q}$ .*

*Dimostrazione.* Sia  $(X, d)$  uno spazio metrico numerabile e  $\mathbb{Q}$  lo spazio dei numeri razionali.

Definiamo una metrica  $\rho$  sull'insieme numerabile  $X \times \mathbb{Q}$  ponendo  $\rho((x, p), (y, q)) = d(x, y) + |p - q|$ .  $X$  è quindi omeomorfo a  $X \times \{0\}$  e  $X \times \mathbb{Q}$  non ha punti isolati.

Per il teorema appena dimostrato  $X \times \mathbb{Q}$  è omeomorfo a  $\mathbb{Q}$ , quindi  $X$  è omeomorfo a un sottospazio di  $\mathbb{Q}$ .  $\square$

## 2. Bibliografia

- [FD] FREDERICK K., DASHIELL JR. (2021), *Countable Metric Spaces Without Isolated Points*, The American Mathematical Monthly, 128:3, 265-267

# I NUMERI NASCOSTI NELLA SCALA MUSICALE

ELISA PEDRINI

## 1. Introduzione

E' noto che la musica e la matematica, sebbene siano due discipline apparentemente molto diverse, abbiano un forte legame. Ma probabilmente pochi saprebbero argomentare tale affermazione. Ci sono diversi collegamenti tra di esse, uno tra questi si ritrova nelle scale musicali.

Per poter procedere in questo discorso è necessaria inizialmente solo la seguente nozione:

DEFINIZIONE 1. La **frequenza di un suono** ( $f$ ) è il numero di vibrazioni in un secondo che danno origine al suono.

La domanda alla base del discorso consiste nel chiedersi quali suoni si possono riprodurre all'interno della gamma delle frequenze udibili, che corrisponde alle frequenze tra la minima  $f_m = 20Hz$  e la massima  $f_M = 20000Hz$ . Se si pensa per esempio al canto o ad alcuni strumenti come il violino o il trombone, sembrerebbe una domanda inutile: ovviamente si possono riprodurre tutte le frequenze nella gamma delle udibili. Ma al contrario la maggior parte degli strumenti consente di riprodurre soltanto un numero finito di suoni, per esempio tutti gli strumenti a tastiera (si pensi al pianoforte con 88 tasti, corrispondenti ciascuno ad una nota), gli strumenti a fiato con i fori o con le chiavi, la chitarra, gli strumenti a percussione,... Questi si dicono strumenti a **suono determinato**.

Il problema in cui ci si riscontra è dunque quello di scegliere un numero finito di frequenze di riferimento, alle quali ogni strumento possa riferirsi, per consentire di avere un linguaggio della musica, denominando tali frequenze con determinati nomi per arrivare ad avere un vero e proprio alfabeto musicale.

In queste pagine verrà dunque affrontata tale questione, ovvero vedremo un percorso verso la costruzione di una **scala musicale**, che consiste nella scelta di  $\{ f_1, f_2, \dots, f_n \}$  frequenze a cui tutti gli strumenti potranno riferirsi ed adeguarsi per eseguire musica.

## 2. Il principio dell'ottava

Il primo passo per la ricerca delle frequenze da scegliere per creare una scala musicale si basa su un principio, presente fin dall'antichità, che consente di semplificare

la selezione dei suoni di riferimento sulla base dell'osservazione di analogie tra di essi:

**PRINCIPIO 1. (Principio dell'ottava)**

*Se una scala contiene il suono  $f$ , allora contiene tutti i suoni della forma*

$$2^k f, k \in \mathbb{Z}$$

**DEFINIZIONE 2.** Data una frequenza  $f$ , si definisce **ottava** l'insieme delle frequenze comprese tra  $2^k f$  e  $2^{k+1} f, k \in \mathbb{Z}$ . Si denota con  $(2^k f, 2^{k+1} f)$

Si osserva che un principio è un enunciato non dimostrato, ma preso per vero come concetto primitivo "autoevidente". E' utile, quindi, dare qualche motivazione per convincersi della validità di tale affermazione e per capirne meglio il significato:

- sperimentalmente si può osservare nel canto che ognuno ha una propria estensione vocale, ovvero una gamma di frequenze che è in grado di produrre (analogo discorso si può fare per i diversi strumenti musicali). Se viene chiesto ad un pubblico di riprodurre un suono, per esempio intonando i 440Hz del diapason, ogni persona (supponendo di avere un pubblico intonato) automaticamente riprodurrà tale nota ma all'interno della propria estensione vocale. Significa che, se si pone attenzione alle note riprodotte, si sentirà lo stesso suono ma più o meno grave a seconda del soggetto. Quello che è successo è che qualcuno avrà riprodotto esattamente la frequenza  $f = 440\text{Hz}$ , altri invece possono aver riprodotto  $2f = 880\text{Hz}$ , ovvero un suono identico ma più acuto, oppure  $\frac{1}{2}f = 220\text{Hz}$ , ovvero un suono più grave.
- una seconda motivazione è di tipo fisiologico: le vibrazioni delle onde sonore entrano nell'orecchio e vengono trasmesse fino all'orecchio interno dove si trova la coclea. Quest'ultima ha una struttura a chiocciola, all'interno della quale si trova un liquido che, muovendosi, consente l'attivazione dei recettori nervosi e quindi la trasmissione dell'impulso al cervello. E' proprio la forma ad elica a dare spiegazione del principio dell'ottava, infatti suoni di differente frequenza mettono in moto nell'orecchio regioni diverse della coclea. La lunghezza di una porzione di elica che corrisponde ad una certa ottava è proporzionale alla lunghezza di quest'ultima (intesa come la distanza euclidea tra due numeri reali, vedi paragrafo 3)

Questo principio permette di semplificare il procedimento verso la costruzione di una scala musicale, perché consente di stabilire i suoni di riferimento soltanto all'interno di un'ottava, poiché richiedere la presenza di un suono  $f$  implicherà automaticamente la presenza di tale suono in tutte le altre ottave tra le frequenze udibili. Si può dunque dare una nuova definizione di ciò che stiamo cercando:

**DEFINIZIONE 3.** Data una frequenza di riferimento  $f_0$ , una **scala musicale** è l'insieme  $\{f_0, f_1, \dots, f_{N-1}, f_N = 2f_0\}$  dato dalla scelta di  $N - 1$  suoni intermedi crescenti nel dominio dell'ottava  $(f_0, 2f_0)$ .



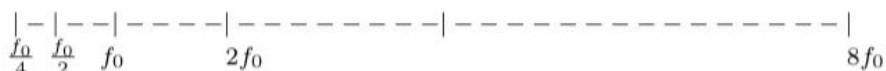
### 3. Congruenza, intervalli e metrica sulla scala musicale

DEFINIZIONE 4. Un suono  $f_i$  si dice **congruente** ad un altro suono  $f_j$  se

$$f_i = 2^k f_j \quad \text{con } k \in \mathbb{Z}$$

OSSERVAZIONE. Tale definizione è una relazione di equivalenza (si verifica banalmente)

Se si pensa alla totalità della gamma delle frequenze udibili come ad un segmento sulla retta reale, fissata una frequenza  $f_0$  e rappresentando alcuni suoni congruenti ad essa, si osserva che la distanza euclidea tra tali frequenze non rimane costante:



D'altra parte, l'identificazione dei suoni deve generare una regola che permetta di ritenere uguali i segmenti con gli estremi identificati. Si può però osservare che ciò che rimane costante è il rapporto tra i valori delle frequenze congruenti successive, ovvero agli estremi di un'ottava:

$$\frac{2^{k+1}}{2^k} = 2 \quad k \in \mathbb{Z}$$

Questa osservazione induce all'introduzione di una distanza tra note, basata appunto sul rapporto:

DEFINIZIONE 5. Date due frequenze  $f_i, f_j$ , si definisce **intervallo musicale** tra  $f_i$  e  $f_j$ , il seguente rapporto:

$$I(f_i, f_j) = \frac{\max\{f_i, f_j\}}{\min\{f_i, f_j\}}$$

OSSERVAZIONE. Si osserva che tale definizione accoglie la richiesta che gli intervalli tra note congruenti siano uguali:

$$I(2^k f_i, 2^k f_j) = \frac{2^k f_i}{2^k f_j} = \frac{f_i}{f_j} = I(f_i, f_j) \quad k \in \mathbb{Z}; f_j \leq f_i$$

Tale definizione si avvicina ad un concetto di distanza tra note, ma non è una vera e propria metrica sull'insieme delle frequenze: banalmente non viene soddisfatta la richiesta che la distanza valga 0 quando  $x = y$ , infatti

$$I(x, x) = \frac{x}{x} = 1 \neq 0$$

Si può però facilmente modificare la definizione per avere una vera e propria distanza dal punto di vista matematico nel seguente modo:

$$d(f_i, f_j) = \log_2(I(f_i, f_j))$$

Da qui in avanti, però, verrà usata soltanto la nozione di intervallo per un discorso di coerenza, perché gran parte del percorso verso la costruzione delle varie scale musicali è di datazione molto antica (si parlerà anche di Pitagora dell'antica Grecia), mentre la nascita del logaritmo avvenne nel 1614 da Nepero.

Si può ora dare una nuova definizione di congruenza di più suoni:

DEFINIZIONE 6. Un insieme ordinato di suoni  $\{f_1, \dots, f_{N-1}, f_N\}$  (non necessariamente appartenenti alla stessa ottava) si dice **congruente** ad un altro insieme ordinato (con lo stesso numero di suoni)  $\{\tilde{f}_1, \dots, \tilde{f}_N\}$  se

$$I(\tilde{f}_1, f_1) = I(\tilde{f}_2, f_2) = \dots = I(\tilde{f}_N, f_N)$$

o equivalentemente se:

$$I(f_2, f_1) = I(\tilde{f}_2, \tilde{f}_1); \quad I(f_3, f_2) = I(\tilde{f}_3, \tilde{f}_2); \quad \dots; \quad I(f_N, f_{N-1}) = I(\tilde{f}_N, \tilde{f}_{N-1})$$

OSSERVAZIONE. Tale definizione è una relazione di equivalenza (si verifica banalmente)

**Proprietà di composizione degli intervalli:** se  $I(f_i, f_j)$  è un intervallo tra frequenze che comprendono il suono  $f_k$  (per esempio  $f_i \leq f_k \leq f_j$ ) allora

$$I(f_i, f_j) = I(f_i, f_k) \cdot I(f_k, f_j)$$

Infatti:  $I(f_i, f_k) \cdot I(f_k, f_j) = \frac{f_k}{f_i} \cdot \frac{f_j}{f_k} = \frac{f_j}{f_i} = I(f_i, f_j)$

#### 4. Vari tipi di scala

Ora che si hanno un po' di nozioni sulle frequenze, si può affrontare la questione centrale: quali sono i criteri con cui si sceglie una scala musicale? (Si ricorda che siamo alla ricerca dell'insieme di frequenze crescenti  $\{f_0, f_1, \dots, f_{N-1}, f_N = 2f_0\}$ , data una frequenza di riferimento  $f_0$ )

Bisogna fare due scelte:

- (1) la scelta del numero  $N$ , ovvero scegliere il numero di note che compongono la scala musicale. Questa è una scelta puramente musicale, dipendente dalle culture e dal genere musicale, vengono ora elencati soltanto alcuni esempi senza affrontare ulteriormente questo aspetto:
  - $N = 5$ : la scala pentatonica, utilizzata in Cina e anche dal Jazz;
  - $N = 6$ : la scala esatonale, utilizzata in Oriente e, per esempio, da Debussy;
  - $N = 7$ : per esempio la scala Blues, la scala Araba e la scala enigmatica;
  - $N = 12$ : per esempio la scala cromatica

E così via...

- (2) Una volta stabilito  $N$ , si devono fissare i gradini  $G_1, G_2, \dots, G_N$ ,

$$G_i = I(f_{i-1}, f_i) \quad i \in \{1, 2, \dots, N\}$$

cioè il valore dei vari intervalli tra le frequenze, ovvero si deve **accordare** la scala. Questo è il punto principale in cui la matematica entra in gioco, nei prossimi paragrafi si vedranno alcune differenti scelte di gradini che danno origine a scale musicali diverse.

La scelta con cui verranno presentate le diverse scale non segue l'ordine cronologico della loro nascita, al contrario: si partirà con la scala più recente, che è quella che si utilizza al giorno d'oggi, per poi poterla confrontare con le precedenti alternative.

## 5. La scala Equabile

Si pone  $N = 12$ , come scegliere i vari gradini? La scelta più spontanea è quella di suddividere l'ottava ( $f_0, 2f_0$ ) in 12 gradini uguali. Ponendo

$$G_1 = G_2 = \dots = G_N = r, \quad r \in \mathbb{R}$$

Si ha che i suoni della scala formano una progressione geometrica di ragione  $r$ , infatti:

$$\frac{f_1}{f_0} = \frac{f_2}{f_1} = \dots = \frac{2f_0}{f_{N-1}} = r$$

Da cui si ricava, usando l'intervallo di un'ottava e il principio di composizione degli intervalli:

$$2 = \frac{2f_0}{f_0} = \frac{2f_0}{f_{N-1}} \cdot \frac{f_{N-1}}{f_{N-2}} \cdot \dots \cdot \frac{f_1}{f_0} = r^N$$

E quindi

$$r = \sqrt[N]{2}$$

Ora si possono calcolare tutti gli intervalli della scala  $I(f_i, f_0)$  con  $i \in \{1, \dots, N\}$  utilizzando sempre il principio di composizione degli intervalli, per esempio:

$$I(f_2, f_0) = r \cdot r = \sqrt[2]{2}$$

e così via, si ottengono

$$I(f_i, f_0) = r \cdot r \cdot \dots \cdot r = \sqrt[i]{2^i} \quad i \in \{1, \dots, N\}$$

OSSERVAZIONE. Tutti questi intervalli sono numeri irrazionali. Questo è il motivo per cui cronologicamente fu l'ultima scala adottata, tanto che la nascita di questo sistema temperato equabile è stato teorizzato nel XVI secolo, con Stevino e poi con Vincenzo Galilei, ma la sua adozione fu molto graduale a causa delle difficoltà nel realizzare l'accordatura degli strumenti: solo nel XIX secolo furono trovati metodi che permisero la realizzazione con precisione dell'accordatura equabile.

Questo temperamento, però, è quello che prevalse sugli altri perché ha un grande vantaggio: la **proprietà di trasposizione**:

Data una melodia, ovvero una successione di note  $\{f_{i_1}, \dots, f_{i_n}\}$ , si può ottenere una melodia congruente semplicemente trasponendo tutti gli indici di una stessa quantità  $k \in \mathbb{Z}$ , cioè prendendo  $\{f_{i_1+k}, \dots, f_{i_n+k}\}$ . All'orecchio risulta esattamente la stessa composizione di note, questo perché nella musica contano gli intervalli tra le note, non la loro effettiva altezza. Questo ha una grande utilità nella pratica, perché consente con facilità di poter adattare agilmente i brani a differenti altezze a seconda delle proprie esigenze e degli strumenti.

La scala musicale così costruita contenente tutte le 13 frequenze è detta **scala cromatica**

L'intervallo tra due frequenze successive in questa scala è detto **semitono** (corrisponde al valore di ogni gradino). L'intervallo  $I(f_i, f_{i+2})$ , dove  $f_i$  è una qualsiasi frequenza della scala cromatica, è detto **tono** (è composto da due semitoni). Esempi di altri intervalli:

- $I(f_i, f_{i+4})$  è detto intervallo di **terza**

- $I(f_i, f_{i+5})$  è detto intervallo di **quarta**
- $I(f_i, f_{i+7})$  è detto intervallo di **quinta**

Il sottoinsieme della scala cromatica costituito dalle frequenze  $\{f_0, f_2, f_4, f_5, f_7, f_9, f_{11}, f_{12}\}$  è detta **scala diatonica maggiore**

OSSERVAZIONE. Prendendo come frequenza di riferimento  $f_0 = \text{DO}$ , se si considera la scala diatonica maggiore si ritrova la famosa scala  $\{\text{DO, RE, MI, FA, SOL, LA, SI, DO}\}$ . Invece l'aggiunta di diesis (#) o bemolle (b) a tali note consente di nominare tutti i suoni della scala cromatica:

$\{\text{DO, DO\#} = \text{REb, RE, RE\#} = \text{MIb, MI, FA, FA\#} = \text{SOLb, SOL, SOL\#} = \text{LAB, LA, LA\#} = \text{SIB, SI, DO}\}$

## 6. La scala Pitagorica

Un'altra scala molto diffusa prima di accettare l'uso dei numeri irrazionali è la **scala pitagorica**, nata dalle osservazioni di Pitagora sulle vibrazioni di una corda: Il principio dell'ottava lo si ritrova confrontando le frequenze della vibrazione di una corda e della corda stessa divisa a metà, ritrovando due suoni identici ma a differenti altezze. Pitagora allora provò a dividere la corda in tre parti uguali e confrontò i suoni ottenuti dalla vibrazione della corda intera e di un suo terzo ottenendo suoni **consonanti**, ovvero suoni che suonati in successione o in simultanea generano una sensazione di riposo e appagamento. Quindi per la costruzione di questa scala si aggiunge un altro principio oltre al principio dell'ottava:

PRINCIPIO 2. (*Principio della quinta*)

Se una scala contiene il suono  $f$ , allora contiene tutti i suoni della forma

$$3^k f, k \in \mathbb{Z}$$

Il nome "principio della quinta" si comprende dalla seguente

OSSERVAZIONE. Preso  $f_0$  come suono di riferimento, per il principio appena citato si ha che il suono  $3f_0$  deve essere incluso nella scala. Chiaramente  $3f_0$  non si trova all'interno dell'ottava  $(f_0, 2f_0)$ , ma basta moltiplicare per  $2^k$ , con  $k \in \mathbb{Z}$  opportuno, per trasportarlo nell'ottava desiderata: in questo caso si ritrova  $\frac{3}{2}f_0 = 1,5f_0$ . Operando un confronto con i suoni della scala equabile si ha che

$$\sqrt[12]{2^7} f_0 \approx 1,49830$$

quindi  $f_7$  è il suono maggiormente vicino a  $\frac{3}{2}f_0$  e lo scarto è poco più che percettibile. Si definisce  $I(f_0, \frac{3}{2}f_0)$  **quinta pitagorica** e così ogni intervallo  $I(f_{i_1}, f_{i_2}) = \frac{3}{2}$  (perché, come si è già visto,  $I(f_i, f_{i+7})$  della scala equabile è detto intervallo di **quinta**)

Analogamente ripetendo lo stesso procedimento si ottiene che

$$I(\frac{3}{2}f_0, 2f_0) = \frac{4}{3} = 1,333... \approx f_5 = 1,33484...$$

e ogni intervallo del valore di  $\frac{4}{3}$  è detto **quarta pitagorica**.

A questo punto si può costruire la scala, partendo dal suono di riferimento  $f_0$  e

osservando che utilizzando il principio dell'ottava e della quinta si devono avere i suoni della forma

$$f = 2^m 3^n f_0 \quad m, n \in \mathbb{Z}$$

OSSERVAZIONE. moltiplicare o dividere per  $2^k$ ,  $k \in \mathbb{N}$  significa portare la nota rispettivamente sopra o sotto di  $k$  ottave. Analogamente moltiplicare o dividere per  $3^k$ ,  $k \in \mathbb{N}$  significa portare la nota rispettivamente sopra o sotto di  $k$  quinte.

PROPOSIZIONE. I numeri  $2^m 3^n f_0$  ( $m, n \in \mathbb{Z}$ ) sono razionali, infiniti e tutti distinti tra loro (se gli esponenti sono distinti)

*Dimostrazione.* Basta dimostrare che sono tutti distinti tra loro, in quanto il resto dell'enunciato è ovvio.

Sia

$$2^{m_1} 3^{n_1} = 2^{m_2} 3^{n_2} \quad m_1, m_2, n_1, n_2 \in \mathbb{Z}$$

si ha

$$2^{m_1 - m_2} = 3^{n_2 - n_1}$$

Se fosse  $m_1 > m_2$  allora anche  $n_2 > n_1$  dato che  $2^{m_1 - m_2} > 1$ . Ma quindi si starebbero uguagliando un numero dispari ed un numero pari. D'altra parte se  $m_1 < m_2$  si conclude allo stesso modo ragionando sugli inversi  $2^{m_2 - m_1}$  e  $3^{n_1 - n_2}$ . Quindi l'unica possibilità è che  $m_1 = m_2$  e  $n_1 = n_2$   $\square$

COROLLARIO 7. Non è possibile determinare  $m \neq 1$  e  $n \neq 0$  in modo tale che  $2^m 3^n$  dia il numero 2.

OSSERVAZIONE. tale corollario è detto "**ciclo delle quinte che non si chiude**": a partire da  $f_0$ , se interviene almeno una quinta ( $n \neq 0$ ) non si potrà mai tornare a  $2f_0$

Per definizione di scala musicale, ora allora è necessario fare una selezione di un numero finito di tali suoni. Per porre un termine all'infinita produzione di suoni, va escogitato un criterio iterativo, da interrompere quando si ritiene opportuno. In questo compito si userà anche il fatto di conoscere già la scala equabile operando direttamente un confronto.

A partire da un suono  $f_0$  si selezionano i suoni tra  $2^m 3^n f_0$  avanzando e retrocedendo per intervalli di quinta pitagorica e riportando eventualmente i suoni nell'ottava di riferimento, se questa viene oltrepassata:

(1) procedendo per quinte si hanno i suoni:

$$\left(\frac{3}{2}\right)^{-m_1} f_0, \dots, \left(\frac{3}{2}\right)^{-1} f_0, f_0, \frac{3}{2} f_0, \left(\frac{3}{2}\right)^2 f_0, \dots, \left(\frac{3}{2}\right)^{m_2} f_0$$

(2) riportando i suoni nell'ottava di riferimento ( $f_0, 2f_0$ ):

$$\text{per } k_1 = 1, \dots, m_1 : f_{h_1, k_1} = 2^{h_1} \left(\frac{3}{2}\right)^{-k_1} f_0, \text{ in modo che } 1 < \frac{f_{h_1, k_1}}{f_0} < 2, \quad h_1 \geq 0$$

$$\text{per } k_2 = 1, \dots, m_2 : f_{h_2, k_2} = 2^{h_2} \left(\frac{3}{2}\right)^{k_2} f_0, \text{ in modo che } 1 < \frac{f_{h_2, k_2}}{f_0} < 2, \quad h_2 \leq 0$$

Con questo procedimento si osserva che dopo 12 passi sia in avanti che all'indietro a partire da  $f_0$ , ovvero con  $k_1 = k_2 = 12$ , i valori ottenuti sono molto prossimi a 1 (all'indietro) e a 2 (in avanti), ovvero si ritrovano suoni molto vicini a  $f_0$  e  $2f_0$  rispettivamente e questo offre la possibilità di chiudere la scala, prevedendo che dal 12esimo passo in poi i calcoli riportino valori quasi coincidenti con quelli già trovati. Nelle prossime due tabelle vengono riportati i valori dei 26 suoni trovati che corrispondono ai suoni della scala pitagorica e vengono posti a confronto con le note della scala equabile.

$f_{h_1, k_1} / f_0 = 2^{h_1} \left(\frac{3}{2}\right)^{k_1}$	$f_{h_2, k_2} / f_0 = 2^{h_2} \left(\frac{3}{2}\right)^{k_2}$
$2 \left(\frac{3}{2}\right)^{-1} = \frac{4}{3} = 1, \bar{3}$	$\frac{3}{2} = 1, 5$
$2^2 \left(\frac{3}{2}\right)^{-2} = \frac{16}{9} = 1, \bar{7}$	$2^{-1} \left(\frac{3}{2}\right)^2 = \frac{9}{8} = 1, 125$
$2^2 \left(\frac{3}{2}\right)^{-3} = \frac{32}{27} = 1, \overline{185}$	$2^{-1} \left(\frac{3}{2}\right)^3 = \frac{27}{16} = 1, 6875$
$2^3 \left(\frac{3}{2}\right)^{-4} = \frac{128}{81} = 1, 58024\dots$	$2^{-2} \left(\frac{3}{2}\right)^4 = \frac{81}{64} = 1, 26562\dots$
$2^3 \left(\frac{3}{2}\right)^{-5} = \frac{256}{243} = 1, 05349\dots$	$2^{-2} \left(\frac{3}{2}\right)^5 = \frac{243}{128} = 1.89843\dots$
$2^4 \left(\frac{3}{2}\right)^{-6} = \frac{1024}{729} = 1, 40466\dots$	$2^{-3} \left(\frac{3}{2}\right)^6 = \frac{729}{512} = 1, 42382\dots$
$2^5 \left(\frac{3}{2}\right)^{-7} = \frac{4096}{2187} = 1, 87288\dots$	$2^{-4} \left(\frac{3}{2}\right)^7 = \frac{2187}{2048} = 1, 06787\dots$
$2^5 \left(\frac{3}{2}\right)^{-8} = \frac{8192}{6561} = 1, 24859\dots$	$2^{-4} \left(\frac{3}{2}\right)^8 = \frac{6561}{4096} = 1, 60180\dots$
$2^6 \left(\frac{3}{2}\right)^{-9} = \frac{32768}{19683} = 1, 66478\dots$	$2^{-5} \left(\frac{3}{2}\right)^9 = \frac{19683}{16384} = 1, 20135\dots$
$2^6 \left(\frac{3}{2}\right)^{-10} = \frac{65536}{59049} = 1, 10985\dots$	$2^{-5} \left(\frac{3}{2}\right)^{10} = \frac{59049}{32768} = 1, 80203\dots$
$2^7 \left(\frac{3}{2}\right)^{-11} = \frac{262144}{177147} = 1, 47981\dots$	$2^{-6} \left(\frac{3}{2}\right)^{11} = \frac{177147}{131072} = 1, 35152\dots$
$2^8 \left(\frac{3}{2}\right)^{-12} = \frac{1048576}{531441} = 1, 97308\dots$	$2^{-7} \left(\frac{3}{2}\right)^{12} = \frac{531441}{524288} = 1, 01364\dots$

scala	E	scala	P
*	1	1	$2^{-7}(\frac{3}{2})^{12}$
	$\sqrt[12]{2} = 1,05946\dots$	$2^3(\frac{3}{2})^{-5}$	$2^{-4}(\frac{3}{2})^7$
*	$\sqrt[12]{2^2} = 1,12246\dots$	$2^{-1}(\frac{3}{2})^2$	$2^6(\frac{3}{2})^{-10}$
	$\sqrt[12]{2^3} = 1,18921\dots$	$2^2(\frac{3}{2})^{-3}$	$2^{-5}(\frac{3}{2})^9$
*	$\sqrt[12]{2^4} = 1,25992\dots$	$2^{-2}(\frac{3}{2})^4$	$2^5(\frac{3}{2})^{-8}$
*	$\sqrt[12]{2^5} = 1,33484\dots$	$2(\frac{3}{2})^{-1}$	$2^{-6}(\frac{3}{2})^{11}$
	$\sqrt[12]{2^6} = 1,41421\dots$	$2^4(\frac{3}{2})^{-6}$	$2^{-3}(\frac{3}{2})^6$
*	$\sqrt[12]{2^7} = 1,49830\dots$	$\frac{3}{2}$	$2^7(\frac{3}{2})^{-11}$
	$\sqrt[12]{2^8} = 1,58740\dots$	$2^3(\frac{3}{2})^{-4}$	$2^{-4}(\frac{3}{2})^8$
*	$\sqrt[12]{2^9} = 1,68180\dots$	$2^{-1}(\frac{3}{2})^3$	$2^6(\frac{3}{2})^{-9}$
	$\sqrt[12]{2^{10}} = 1,78180\dots$	$2^2(\frac{3}{2})^{-2}$	$2^{-5}(\frac{3}{2})^{10}$
*	$\sqrt[12]{2^{11}} = 1,88775\dots$	$2^{-2}(\frac{3}{2})^5$	$2^5(\frac{3}{2})^{-7}$
*	2	2	$2^8(\frac{3}{2})^{-12}$

OSSERVAZIONE. Si osserva che i 26 suoni si dispongono a coppie offrendo un'approssimazione per difetto ed una per eccesso delle frequenze irrazionali della scala equabile. Dal punto di vista matematico tale processo è soddisfacente: viene sostanzialmente riprodotta la scala equabile. Dal punto di vista musicale però si è deciso di fare un'ulteriore selezione tra questi suoni ottenuti: i suoni della scala diatonica maggiore vengono presi solamente una volta (il primo valore incontrato nel procedimento), mentre gli altri vengono presi entrambi e vengono assegnati con il nome "bemolle (b)" per quelli che approssimano per difetto e "diesis (#)" per quelli che approssimano per eccesso.

Quindi in questa scala si ha una maggiore ricchezza di suoni, poiché diesis e bemolli non coincidono. Lo svantaggio di questo temperamento è il fatto che la scala ha una forte dipendenza dal valore  $f_0$  di riferimento poiché, a parità di passaggi, una diversa frequenza di partenza dà origine a suoni differenti.

E' chiaro anche quindi che in tale contesto non vale la proprietà di trasposizione!

## 7. La scala naturale

La costruzione di questa scala segue inizialmente la stessa linea di quella pitagorica, osservando che la divisione della corda in 2, 3, 4(= 2<sup>2</sup>) parti uguali sono già state prese in considerazione dal principio dell'ottava e della quinta. L'idea è, quindi, di dividere la corda in 5 parti uguali, divisione che dà origine ancora una volta ad un suono consonante rispetto a quello prodotto dalla corda intera. Si aggiunge così un terzo principio:

**PRINCIPIO 3. (Principio della terza)**

Se una scala contiene il suono  $f$ , allora contiene tutti i suoni della forma

$$5^k f, k \in \mathbb{Z}$$

OSSERVAZIONE. Il nome "principio della terza" viene spiegato in maniera analoga a quanto visto per la scala pitagorica:  $5f_0$  riportato nell'ottava corrisponde a  $\frac{5}{4}f_0 = 1,25f_0$  che è molto prossimo a  $\sqrt[12]{2^4}f_0$ . Ogni intervallo di valore  $\frac{5}{4}$  è denominato così **terza maggiore naturale**.

A questo punto, esattamente come precedentemente, bisogna fare una selezione tra gli infiniti suoni distinti e dati dai numeri razionali

$$f = 2^m 3^n 5^p f_0 \quad m, n, p \in \mathbb{Z}$$

Si potrebbe pensare di procedere per quinte e terze, ma in questo caso l'incasellatura tra i vari suoni appare molto più articolata. Si predilige allora una costruzione differente, geometrica, basata sulla divisione armonica. I dettagli di tale costruzione purtroppo non verranno affrontati, si è voluto comunque citare tale scala per poter operare un confronto tra le tre scale e poter così osservare che quest'ultima ha frequenze molto più semplici. Quindi nel confronto tra la scala pitagorica e quella naturale si ha: un procedimento di costruzione più immediato nella prima scala ma un risultato migliore come semplicità di frequenze nella seconda.

	E	P	N
DO	1	1	1
RE	$\sqrt[12]{2^2} = 1,12246\dots$	$2^{-3}3^2 = 1,125$	$2^{-3}3^2 = 1,125$
MI	$\sqrt[12]{2^4} = 1,25992\dots$	$2^{-6}3^4 = 1,26562\dots$	$2^{-2}5 = 1,25$
FA	$\sqrt[12]{2^5} = 1,33484\dots$	$2^23^{-1} = 1,\bar{3}$	$2^23^{-1} = 1,\bar{3}$
SOL	$\sqrt[12]{2^7} = 1,49830\dots$	$2^{-1}3 = 1,5$	$2^{-1}3 = 1,5$
LA	$\sqrt[12]{2^9} = 1,68180\dots$	$2^{-4}3^3 = 1,6875$	$3^{-1}5 = 1,\bar{6}$
SI	$\sqrt[12]{2^{11}} = 1,88775\dots$	$2^{-7}3^5 = 1,89843\dots$	$2^{-3}3 \times 5 = 1,875$
DO	2	2	2

Ribaltando uno dei vantaggi del sistema temperato equabile, ovvero che gli strumenti ad intonazione fissa suonano ugualmente bene in tutte le tonalità, si può dire che un vantaggio della scala naturale è che, preso qualsiasi intervallo, esiste sempre una tonalità (ovvero una scelta differente della frequenza  $f_0$  di riferimento da cui si costruisce la scala) in cui tale intervallo risulta perfettamente consonante. Tale proprietà non è valida nel temperamento equabile dove un intervallo dissonante (cioè non consonante) rimane tale in qualsiasi tonalità, per la proprietà di trasposizione.

In assenza di precisi metodi di misura delle frequenze, una varietà di differenti temperamenti era possibile, almeno fino al 1831, quando fu introdotto l'uso del diapason per accordatura. E' interessante pensare come sia stato difficile arrivare alla



scala musicale di uso ormai comune e riflettere sul fatto che effettivamente alla base di questo discorso ci sia la matematica e la difficile accettazione dei numeri irrazionali.

### Bibliografia

- [T] F.TALAMUCCI, *La costruzione di una scala musicale attraverso i numeri*  
[F] [http://fisicaondemusica.unimore.it/Temperamento\\_equabile](http://fisicaondemusica.unimore.it/Temperamento_equabile).



# MUSICA E MATEMATICA: UN LEGAME INDISSOLUBILE

EDWARD KEVIN ARANA MEDINA

## 1. Introduzione

*Perché due suoni generano un suono piacevole? esiste un criterio per stabilire se due suoni siano armoniosi?* questi quesiti ci condurranno inaspettatamente verso un risultato fondamentale della teoria della misura, che è una branca dell'analisi reale e complessa, nata a cavallo tra il XIX e il XX secolo, che ha a che fare con gli integrali e la probabilità. Infine utilizzando il cubo di Rubik, gli accordi musicali e un pizzico di teoria dei gruppi troveremo un modo per far convergere i suoni cacofonici a dei suoni piacevoli.

## 2. Alla ricerca di suoni armoniosi

### 2.1. Il problema

Consideriamo due frequenze 220 Hz (che corrisponde alla nota do centrale) e  $r \cdot 220$  Hz,  $r \in \mathbb{R}$ ,  $1 < r <$ , e ci chiediamo per quali valori di  $r$  suonano bene insieme. Pitagora di Samo sicuramente avrebbe risposto dicendo che due frequenze suonano bene insieme se  $r \in \mathbb{Q} \cap (1,2)$ . Inizialmente potremmo essere tratti in inganno a pensare che la risposta di Pitagora sia giusta, infatti per  $r = \frac{3}{2}$  ottengo un intervallo di quinta, per  $r = \frac{4}{3}$  un intervallo di quarta, per  $r = \frac{8}{5}$  un intervallo di sesta maggiore, etc. ma in realtà la maggior parte dei numeri razionali produce suoni cacofonici, per esempio se prendiamo  $r = \frac{211}{198}$  oppure  $r = \frac{1093}{826}$

### 2.2. Perché certe frequenze generano dei suoni piacevoli e altre dei suoni cacofonici?

Una nota musicale è costituita da battiti suonati in rapida successione, per esempio la nota do centrale che corrisponde alla frequenza 220 Hz possiamo pensarla come ad una successione rapida di 220 battiti al secondo. E quando il rapporto tra due frequenze è razionale ( o equivalentemente quando  $r$  è un numero razionale ) si può percepire uno schema intuibile per questi battiti che se rallentato si può percepire come "ritmo" invece che come un armonia. Quando i nostri cervelli riescono a riconoscere questo schema ritmico allora vengono percepite dal nostro cervello come suono piacevole.

OSSERVAZIONE 1. Le frazioni "semplici" ,ossia con numeratore e denominatore che assumono valori non troppo distanti da zero, sembrano suonare bene, mentre le

frazioni "complicate", ovvero con numeratore e denominatore che assumono valori molto distanti da zero, suonano male. E se considero delle frazioni con denominatori bassi ( $< 10$ ), suonerebbero bene.

### 2.3. Ultima accortezza..

Possiamo affermare che le frequenze 220 Hz e  $r \cdot 220$  Hz suonano bene insieme  $\iff r \in \mathbb{Q} \cap (1,2)$ ? No, perché se consideriamo  $r = \frac{e^{\pi-15\pi}}{15}$  notiamo che le frequenze 220 Hz e  $r \cdot 220$  Hz producono un suono gradevole. dunque se  $r \in \mathbb{Q} \cap (1,2)$  non riesco a catturare tutti i suoni armoniosi. Osservo che  $r = \frac{e^{\pi-15\pi}}{15} \rightarrow \frac{4}{3}$ . Ma allora anche i suoni irrazionali suonano bene? Infatti svariati strumenti musicali come il pianoforte non sono accordati per intervalli razionali ma in modo tale che ogni mezzo aumento di tono (semitono) corrisponda al moltiplicare la frequenza per originale per la radice dodicesima di 2, che è irrazionale. Dunque se considero un intervallo musicale, ad esempio una quinta, il rapporto delle loro frequenze non sarà  $\frac{3}{2}$  ma bensì  $2^{(\frac{7}{12})}$  che è irrazionale, ma è molto vicino a  $\frac{3}{2}$  che è un numero razionale. In realtà quello che succede è che tutti le potenze della radice dodicesima di 2 si approssimano con un errore meno del 1 % ad un numero razionale.

TEOREMA 2. le frequenze 220 Hz e  $r \cdot 220$  Hz suonano bene insieme  $\iff r \in \mathbb{R} \cap (1,2)$ ,  $r$  molto vicino a un numero razionale con un denominatore molto basso

### 2.4. Domanda interessante

Supponiamo che ci sia una persona in grado di percepire armoniosi tutte le coppie di note le cui frequenze abbiano un rapporto razionale (anche quelle complicate che noi troveremmo cacofonici). Il quesito che ci poniamo è il seguente: *Questa persona troverebbe tutti i rapporti armoniosi? scegliendo come  $r$  ogni numero reale tra 1 e 2?*

## 3. Teoria della misura

### 3.1. Un problema interessante

TEOREMA 3.  $\exists I_n \in \mathbb{R}$  intervalli aperti tali che :

$$(1) \bigcup_{n=1}^{\infty} I_n \supseteq \mathbb{Q} \cap (0,1)$$

$$(2) \sum_{n=1}^{\infty} |I_n| < 1$$

*Dimostrazione.* Innanzitutto elenchiamo i numeri razionali tra 0 e 1, per farlo elenchiamo prima tutte le frazioni ridotte tra 0 e 1 con denominatore 2, successivamente tutte le frazioni ridotte con denominatore 3, e così via..

$\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}$ , ecc.. In questo modo siamo riusciti ad elencare i numeri razionali tra 0 e 1. Ora dobbiamo assegnare un opportuno intervallo ad ogni razionale. Per fare ciò, ci conviene scegliere intervalli con ampiezze diverse e che siano decrescenti. Siccome abbiamo a che fare con un numero infinito di intervalli possiamo considerare una serie a termini positivi convergente a 1, come per esempio:  $\sum_{n=1}^{\infty} \frac{1}{2^n} = 1$ . Successivamente scegliamo un  $\varepsilon$  tale che  $0 < \varepsilon < 1$  e abbiamo finito, infatti:  $\varepsilon \cdot (\sum_{n=1}^{\infty} \frac{1}{2^n}) = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$ . Ora ci basta assegnare le ampiezze degli intervalli

ai numeri razionali e abbiamo finito, per farlo basta assegnare al n-esimo numero razionale un'ampiezza data dal n-esimo termine della somma. Dunque assegno ad  $\frac{1}{2}$  la palla (o interno circolare) di centro  $\frac{1}{2}$  e di raggio  $\frac{\varepsilon}{4}$ , ad  $\frac{1}{3}$  la palla di centro  $\frac{1}{3}$  e di raggio  $\frac{\varepsilon}{8}$ , e così via.. così facendo la somma delle ampiezze può essere arbitrariamente piccola!  $\square$

OSSERVAZIONE 4. Se consideriamo  $\varepsilon = 0.3$  e scelgo un numero a caso tra 0 e 1, c'è il 70 % di probabilità che si trovi fuori da tutti gli intervalli. Ed un esempio di numero che si trova al di fuori di tutti questi intervalli è  $\frac{\sqrt{2}}{2}$ . L'intervallo del n-esimo numero razionale della lista è molto piccola se prendiamo un n molto grande, e i numeri razionali di quel tipo hanno denominatori molto grandi. Dunque possiamo dedurre che  $\frac{\sqrt{2}}{2}$  non coperto dagli intervalli  $\Rightarrow$  i razionali che sono vicini a  $\frac{\sqrt{2}}{2}$  hanno un denominatore molto grande ma  $\Rightarrow \frac{\sqrt{2}}{2}$  è cacofonica!

Ora consideriamo  $\varepsilon = 0.01$  e consideriamo l'intervallo (1,2) anziché l'intervallo (0,1). Se scelgo un numero a caso tra 0 e 1, c'è il 99 % di probabilità che si trovi fuori da tutti gli intervalli. La maggior parte dei numeri coperti dagli intervalli sono armoniosi.

ESEMPIO 5.  $2^{\frac{7}{12}}$  è davvero molto vicino a  $\frac{3}{2}$ , e cade nel intervallo assegnato a  $\frac{3}{2}$

OSSERVAZIONE 6. Ma quali numeri del 1 % sono cacofonici? Quelli con denominatori elevati e i numeri reali vicinissimi a essi. La persona che era in grado di percepire armoniosi tutte le coppie di note le cui frequenze abbiano un rapporto razionale...trova armonioso precisamente l'1 % dei numeri coperti dagli intervalli. Dunque per quella persona i numeri armoniosi sono rari!

#### 4. Teoria dei gruppi

DEFINIZIONE 7. Un gruppo è un insieme G dotato di un'operazione binaria  $\diamond$  interna tale che vengono soddisfatte le seguenti proprietà:

- (1) Associatività:  $\forall x, y, z \in G$  t.c.  $(x \diamond y) \diamond z = x \diamond (y \diamond z)$
- (2) Esistenza dell'elemento neutro:  $\exists e \in G$  t.c.  $\forall x \in G$   $x \diamond e = e \diamond x = x$
- (3) Esistenza dell'inverso:  $\forall x \in G \exists x^{-1} \in G$  t.c.  $x \diamond x^{-1} = x^{-1} \diamond x = e$

##### 4.1. Cubo di Rubik e la teoria dei gruppi

Un cubo di Rubik possiamo vederlo come un gruppo dove gli elementi del gruppo sono tutte le possibili configurazioni delle facce del cubo, e l'operazione è la rotazione delle facce del cubo. La cardinalità del gruppo è 43.252.003.274.489.856.000 dunque non è una buona idea cercare di risolvere un cubo di Rubik casualmente. Ma come risolvo il cubo di Rubik? Usando la teoria dei gruppi possiamo analizzare il cubo e determinare una sequenza di permutazioni che permette di risolvere il cubo.

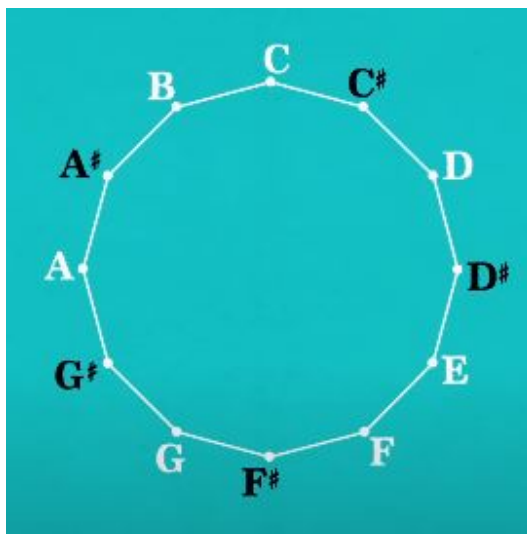
ESEMPIO 8. Indichiamo con  $a^3$  la rotazione di  $90^\circ$  di una faccia del cubo. Posso rappresentare una permutazione che risolve il cubo di Rubik in questo modo:  $a^3$ ("verde") +  $a^3$ ("rosso") +  $a^3$ ("bianco"), ovvero ruoto di  $90^\circ$  la faccia di color verde, poi la faccia

rossa e infine quella bianca, in questo modo ho risolto il cubo, con una sequenza di permutazioni.

#### 4.2. Parliamo di accordi..

DEFINIZIONE 9. Un accordo in musica è l'emissione simultanea di tre o più note.

4.2.1. *Come creare le triadi (accordi composte da sole tre note).* Scriviamo le 12 note musicali lungo una circonferenza dove il polo nord corrisponde alla nota DO. La prima nota la scegliete casualmente, l'altra la prendete di quattro posti in senso orario, e l'ultima a tre posti in senso orario.



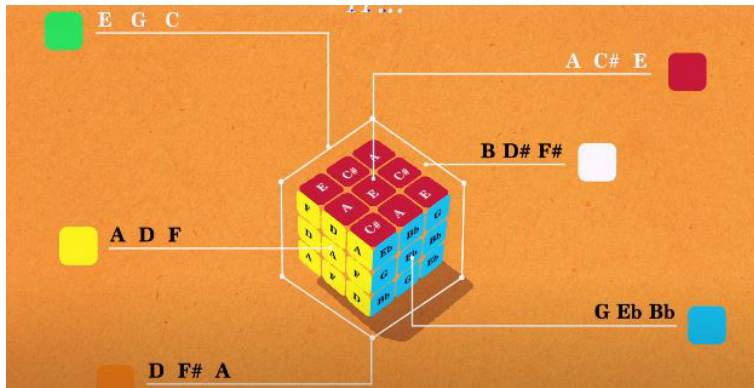
ESEMPIO 10. Do - Mi - Sol, Re - Fa - La, Do - Re - Fa - La

#### 4.3. Convergenza dei suoni cacofonici a suoni armoniosi

Consideriamo un cubo di Rubik risolto e 6 triadi qualsiasi, e assegniamo ad ogni faccia un accordo differente. Ed assegniamo ad ogni "quadrato" della faccia un nome corrispondente ad una nota del accordo, tale che 2 quadratini adiacenti non abbiano lo stesso nome. Rimescolando il cubo di Rubik, ogni faccia non rappresenterebbe più un accordo ma un suono cacofonico..

#### 4.4. Risoluzione del cubo di Rubik

Ora siamo in grado di "aggiustare" i suoni cacofonici e trasformarli in suoni armoniosi, per farlo basta risolvere il cubo di Rubik, infatti la soluzione del cubo posso ovviamente esprimerla come una progressione di suoni che ,gradualmente , parte da un suono disarmonico e si avvicina sempre più ad un suono piacevole!



### Bibliografia

- [M] ELI MAOR (2018), *La musica dai numeri: musica e matematica, da Pitagora a Schoenberg*. Codice edizioni.
- [FM] NICOLA FUSCO, PAOLO MARCELLINI E CARLO SBORDONE (2020), *Lezioni di analisi matematica due*, Zanichelli.
- [S] MICHEAL STAFF (2016), *Group theory 101: How to play a Rubik's Cube like a piano*, Youtube.





# CHE CHAOS!

URMILA BOSISIO

Tutti abbiamo sentito almeno una volta l'espressione "battito d'ali di una farfalla" in relazione a un fenomeno minuscolo che può dare origine a effetti catastrofici; è con questa suggestiva immagine che ci si può addentrare nello studio della teoria del chaos, che si occupa per l'appunto di tutti quei modelli che sulla carta non danno problemi ma, nel momento in cui devono essere impiegati per approssimare il fenomeno che descrivono, utilizzando perciò dei dati iniziali imprecisi dovuti a delle stime o alla precisione macchina del computer, risultano decisamente inefficaci, dando risultati molto diversi da quelli che effettivamente sarebbero previsti dal modello per dei dati teoricamente esatti.

La domanda fondamentale da porsi è quindi: avendo la sicurezza di star approssimando "bene" un dato iniziale in un sistema dinamico, si ha anche una buona approssimazione della soluzione, in generale? La risposta, che si può intuire essere negativa, coinvolge per l'appunto i cosiddetti *sistemi caotici*.

## La nascita della teoria del chaos

È nel XVI secolo che Keplero formula le sue leggi che descrivono le orbite dei pianeti del sistema solare, ed è con Newton che diventano rilevanti e si stabiliscono come universalmente accettate. Il principio su cui si basano è che il moto di due corpi celesti che si attraggono vicendevolmente può essere descritto completamente, senza margine di errore. Nel corso del XIX secolo, il matematico Henri Poincaré si interessa a questo mondo; in particolare, studia il cosiddetto problema dei tre corpi, che sostanzialmente prevede tre corpi celesti interagenti. Può sembrare strano ma aggiungendo un altro corpo vengono coinvolte così tante variabili che tale problema risulta estremamente complesso. In ogni caso, Poincaré dapprima fornisce una soluzione parziale ottenuta tramite approssimazioni, per poi rendersi conto che, con una piccolissima variazione nelle condizioni iniziali del problema, le orbite prodotte risultano completamente diverse.

Per visualizzare questa cosa, si può considerare il pendolo doppio, costituito da due pendoli uno agganciato all'altro. Partendo da una posizione in cui il primo pendolo forma un angolo retto con la verticale e il secondo con l'orizzontale, facendo diverse simulazioni si possono notare orbite completamente diverse.

Questa intuizione di chaos inizia da Poincaré, però ha la sua massima espressione

con l'avvento del calcolatore. In particolare, negli anni Sessanta del secolo scorso un meteorologo dell'MIT chiamato Ed Lorenz sviluppa un programma a computer per risolvere delle equazioni matematiche che possano descrivere, per l'appunto, le condizioni meteorologiche. Il modello:

$$\begin{cases} \dot{x} &= \sigma(y - x) \\ \dot{y} &= x(\rho - z) - y \\ \dot{z} &= xy - \beta z \end{cases} \quad \text{dove } \sigma = 10, \beta = 8/3, \rho = 28$$

risulta non lineare, ma non in modo drammatico. La cosa sorprendente è che, a distanza di poco tempo, rimettendo nel programma gli stessi dati ma con un errore di approssimazione con ordine di grandezza leggermente più grande i risultati ottenuti sono decisamente diversi.

Questo fenomeno si può visualizzare studiando un minimo questo sistema. Si scopre infatti che il sistema ammette due punti di equilibrio che risultano instabili, anche se vi è una direzione stabile. In particolare, si possono immaginare in tre dimensioni due piani su cui giacciono questi due punti di equilibrio, su cui la soluzione tende a spiraleggiare sempre più lontano come un fuoco instabile, e poi si ha una direzione incidente al piano che per l'appunto richiama l'orbita a sé. In totale, si ha un comportamento sorprendente: quello che si può vedere in figura 1; questo oggetto che "attrae" orbite viene definito *attrattore strano*. È facile vedere che l'orbita è molto complessa, e ci si può quindi aspettare che prendendo un punto molto vicino al punto iniziale di questa curva si avrà un risultato diverso, magari un punto starà da una parte a un certo tempo e un altro sull'altro "piano".

### Diverse definizioni

Per dare una forma a questo fenomeno, ci si può addentrare nel caso discreto; qui, la dinamica è in generale descritta da un'iterazione del tipo  $x_{n+1} = f(x_n)$  con  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Come nozione fondamentale, si definisce *punto periodico* di periodo  $N$  un  $\bar{x}$  tale per cui  $\exists N \in \mathbb{N}$  tale che  $f^N(\bar{x}) = \bar{x}$ .

È nel 1975 che Li e Yorke danno la prima definizione di caoticità.

DEFINIZIONE 1. Una mappa  $f: \mathbb{R} \rightarrow \mathbb{R}$  è detta **caotica** (nel senso di Li e Yorke) se possiede uno *scrambled set* non numerabile, cioè un insieme  $S \subset \mathbb{R}$  tale che:

- (1)  $\forall x, y \in S$ , con  $x \neq y$  si ha  $\limsup_{k \rightarrow +\infty} |f^k(x) - f^k(y)| > 0$ ;
- (2)  $\forall x, y \in S$  si ha  $\liminf_{k \rightarrow +\infty} |f^k(x) - f^k(y)| = 0$ ;

per cui vale anche

- (a)  $\forall x \in S, \forall y$  punto periodico con  $x \neq y$  si ha  $\limsup_{k \rightarrow +\infty} |f^k(x) - f^k(y)| > 0$ .

La definizione prevede quindi l'esistenza di un insieme che non contenga punti periodici né definitivamente periodici; non ammetta orbite asintoticamente periodiche; sia non numerabile. In particolare, quindi, coinvolge principalmente concetti metrici.

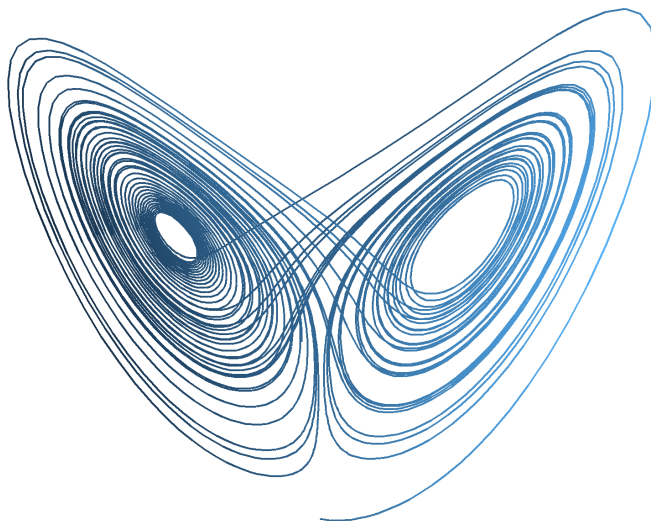


FIGURA 1. Attrattore strano

È nel 1989 che Devaney pubblica *An Introduction to Chaotic Dynamical Systems*, in cui introduce una nuova definizione di caoticità:

DEFINIZIONE 2. Una mappa  $f : I \rightarrow I$  con  $I \subset \mathbb{R}$  intervallo è detta **caotica** (nel senso di Devaney) se:

- (1) è *topologicamente transitiva*, cioè  $\forall U, V$  aperti  $\exists n \in \mathbb{N}$  tali che  $f^n(U) \cap V \neq \emptyset$ ;
- (2) ha dipendenza sensibile dai dati iniziali, cioè  $\exists \delta > 0$  tale che  $\forall x \in I, \forall U$  intorno aperto di  $x, \exists y \in U \setminus \{x\}, \exists n \in \mathbb{N}$  tali che  $|f^n(x) - f^n(y)| > \delta$ ;
- (3) ha un insieme denso di orbite periodiche.

Questa seconda definizione prevede quindi, immaginando di lavorare su un intervallo  $I$  compatto: l'esistenza di un'orbita densa in  $I$ ; per ogni punto, l'esistenza di almeno un altro punto arbitrariamente vicino che viene separato dalla dinamica di almeno  $\delta$  dopo un certo numero di iterate della mappa  $f$ ; una quantità non numerabile di orbite periodiche. Si presti attenzione a quest'ultima condizione: dà infatti un "tocco di regolarità" inaspettato, formalizzando quindi il concetto per cui affinché vi sia chaos sia necessaria la presenza di un certo equilibrio.

Le due definizioni analizzate non sono a sé stanti e, in effetti, sono legate tra loro; in particolare, è possibile dimostrare che la definizione di caoticità data da Devaney risulta più forte di quella proposta da Li e Yorke, ed è per questo motivo quella più utilizzata.

## Regolarità nel chaos

Si può quindi approfondire questo aspetto di regolarità che costituisce una parte importante della definizione di caoticità. Per fare ciò, si possono considerare alcuni esempi fondamentali.

Negli anni '70 si iniziano a studiare modelli discreti per l'evoluzione della popolazione, e nel 1975 in particolare Robert May propone il modello logistico:

$$x_{n+1} = rx_n(1 - x_n) \quad \text{con } r > 0 \text{ tasso di crescita.}$$

Visualizzando la mappa  $f(x) = rx(1 - x)$  definita su  $I = [0, 1]$ , è facile notare che per  $r < 4$  l'iterata di  $f$  sta ancora in  $I$ , da cui si può considerare  $0 < r < 4$ ). Studiando i punti fissi di  $f$  e le sue iterate si nota che, aumentando  $r$ , i punti fissi di  $f$  si "sdoppiano" in punti fissi per  $f^2 = f \circ f$ , e così via, provocando le cosiddette *biforcazioni*: in particolare, si vengono a creare punti periodici con periodo potenze di due, fenomeno che prende il nome di "cascata di biforcazioni con raddoppio di periodo" (si veda figura 2). Si ha in effetti, da un certo valore di  $r$  in poi, si ha un'infinità non numerabile di orbite periodiche. Questo, insieme alle altre due proprietà fondamentali (dipendenza sensibile dai dati iniziali e transitività topologica), garantiscono la caoticità di questa mappa.

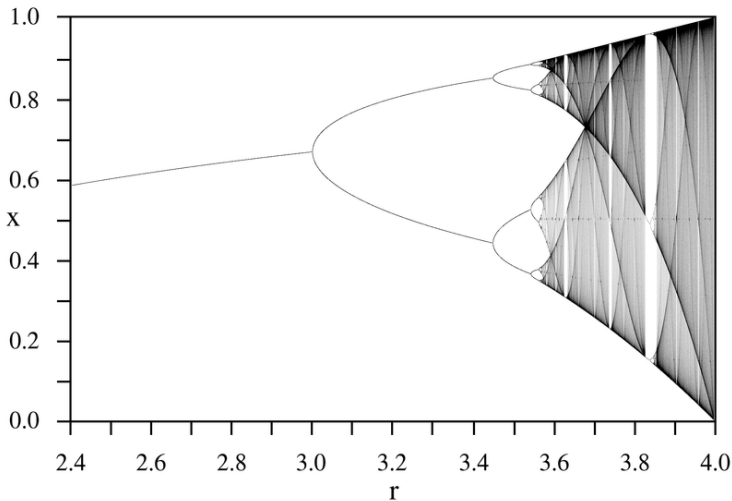


FIGURA 2. Biforcazioni con raddoppio di periodo

Un altro esempio è la mappa del gatto di Arnold, proposta da quest'ultimo matematico, e il cui nome particolare verrà presto spiegato. L'iterazione si basa su una mappa definita sul toro, e in particolare si può considerare definita sul solito intervallo  $I = [0, 1]$  come:

$$\Gamma : (x, y) \rightarrow (2x + y, x + y) \pmod{1}.$$

Studiandola è possibile notare l'esistenza di una direzione di espansione e una di contrazione, la cui azione si può visualizzare in figura 3(a); in particolare, si può considerare la mappa come la distorsione di un'immagine quadrata e la relativa ricomposizione come nello schema sottostante.

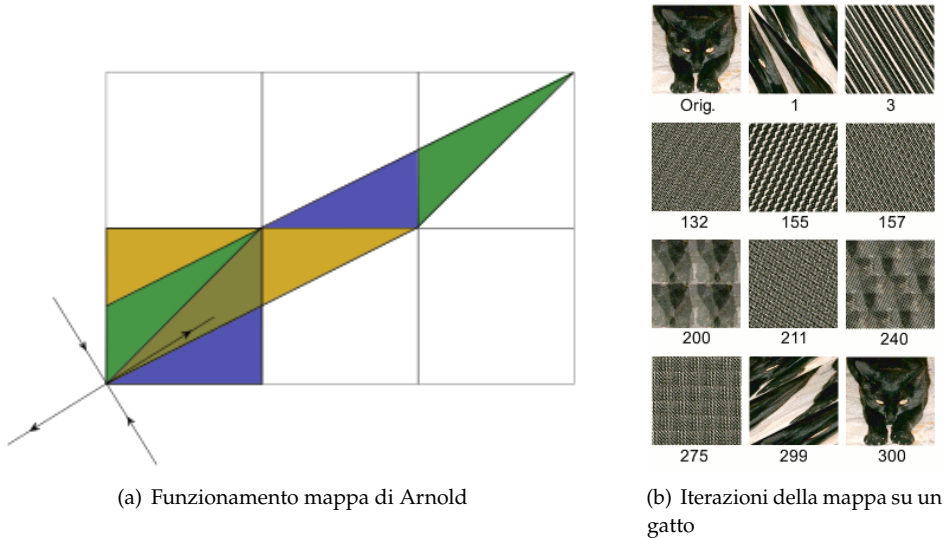


FIGURA 3. Mappa di Arnold

È facile intuire che questa mappa risulta caotica. In particolare, portando avanti il parallelismo con le immagini, si può desumere che deformando e ricomponendo più volte immagini leggermente diverse in questo modo, esse ne risultino spaventosamente diverse. Un aspetto curioso sta nella presenza di punti periodici, e quindi di immagini che vengono, dopo un certo numero di iterazioni, ricostruite. In effetti, il nome della mappa deriva dal fatto che, applicando la mappa su un'immagine stilizzata di un gatto, questa si ripresenti uguale dopo un certo numero di iterazioni. Qui si riporta come esempio l'immagine di un gatto vero (non l'immagine utilizzata da Arnold, ovviamente), ma il risultato è analogo.

### Applicazioni

Ci si potrebbe quindi chiedere quale sia l'effettiva utilità della teoria del caos, e se questi sistemi caotici siano effettivamente circoscritti a degli esempi simpatici come quelli sopra riportati o se esistono fenomeni che possono essere effettivamente studiati a partire da quest'ottica.

È possibile dare quindi una risposta considerando la cosiddetta *circle map*:

$$\theta_{n+1} = \theta_n + \Omega + \varepsilon \sin(\theta_n) \quad \text{mod } 2\pi$$

Tale mappa, ovviamente caotica per certe scelte di valori per  $\Omega$  e  $\varepsilon$ , si può ritrovare in svariati contesti: in fenomeni di oscillazione (ad esempio in elettrocardiogrammi e

nella concentrazione di determinate sostanze in processi biologici, come la diffusione di cellule tumorali nel corpo umano); nell'inarmonicità di strumenti musicali; nella rotazione sincrona nell'orbita dei satelliti.

La conoscenza della caoticità di questa mappa permette quindi, tramite strumenti anche probabilistico-statistici, di controllare la crescita di tumori e di studiare i ritmi cardiaci per prevenire aritmie nei pazienti che ne possono essere soggetti.

Un'applicazione invece più teorica e fisica è quella della nascita della teoria KAM, che prende il nome da Kolmogorov-Arnold-Moser. Tale teoria ha enorme importanza per la fisica classica e trova applicazioni nella fisica del plasma e nella meccanica celeste.

### Bibliografia

- [RM] ROBERT M. MAY (1976), *Simple Mathematical Models with Very Complicated Dynamics*, Nature Vol. 261, pp. 459-466.
- [TY] TIEN-YIEN LI, JAMES A. YORKE (1975), *Period Three Implies Chaos*, The American Mathematical Monthly, Vol.82, pp. 985-992.
- [MF] MITCHELL J. FEIGENBAUM (1975) *Universality in complex discrete dynamic*, Los Alamos Theoretical Division Annual Report
- [D] ROBERT L. DEVANEY (1986), *An Introduction to Chaotic Dynamical Systems*, Benjamin/Cummings, pp. 49-66.
- [KH] ARUNACHALAM KUMAR, B.M. HEGD (2012), *Chaos theory: impact on and applications in medicine*, Journal of Health and Allied Sciences NU, Vol. 2, pp. 93-99.

# COSA C'ENTRA IL CALCOLO DELLE VARIAZIONI CON L'OTTICA GEOMETRICA?

ALESSIA SEVERGNINI

Ce lo dicono Eulero, Lagrange, Hamilton e ancora Maupertuis, che riformulando il Principio di Fermat sull'ottica geometrica in modi quasi equivalenti, hanno scoperto che in quel quasi può nascondersi il seme di una nuova teoria matematica.

Storicamente, i principi variazionali, che qui studiamo nelle formulazioni di Hamilton e Lagrange, divennero un argomento centrale di ricerca nella comunità scientifica un secolo prima di Hamilton, volendo se ne potrebbe assegnare un anno ufficiale di nascita, il 1744, vediamo perché. L'idea in realtà nasce qualche tempo prima, quando la formulazione che Maupertuis aveva dato del principio di Fermat per i raggi dell'ottica geometrica in ambito puramente meccanico suscitò grandi discussioni. La teoria matematica tuttavia nasce con l'idea di Eulero, che ha affrontato e poi risolto un problema più generale in un fondamentale lavoro, pubblicato appunto nel 1744, dal titolo "Methodus inveniendi lineas curvas maximi minimive proprietates gaudentes". Tale studio, in effetti, ha come obiettivo quello di caratterizzare come soluzioni di equazioni differenziali le funzioni che hanno la proprietà di essere punti estremali per un assegnato funzionale di tipo integrale.

L'approccio dei principi variazionali è quindi quello di caratterizzare un movimento o una traiettoria non attraverso un'equazione differenziale, bensì mediante una proprietà di minimo o di massimo rispetto a una famiglia di movimenti o traiettorie.

## La legge di rifrazione

Sono dati due mezzi omogenei 1, 2, separati dalla retta  $r$ , nei quali la luce si propaga rispettivamente con velocità  $v_1, v_2$ . Studio il fascio di luce nel suo percorso da  $A$  a  $B$ , punti posti rispettivamente nei mezzi 1 e 2. La legge di rifrazione afferma che il fascio di luce percorre la curva spezzata  $ACB$  soddisfacente

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2}, \quad (1)$$

dove  $C$  è un punto sulla retta  $r$  tale per cui i segmenti  $AC$  e  $CB$  formano con la verticale rispettivamente angoli  $\theta_1$  e  $\theta_2$ .

Proviamo ora a dare una dimostrazione della legge utilizzando il calcolo delle variazioni.

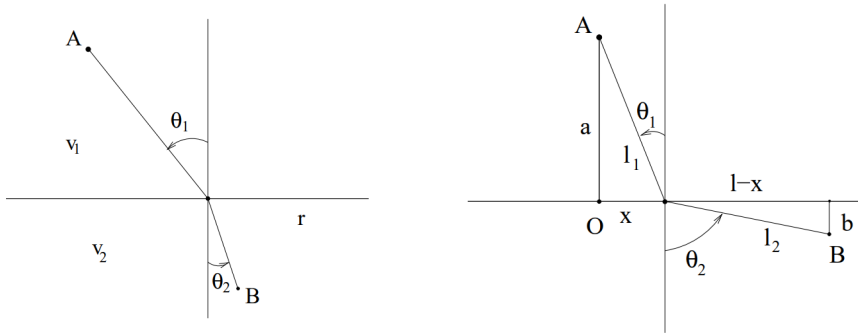


FIGURA 1. Legge di rifrazione

In particolare mi chiedo quale sia la curva che congiunge A e B per cui il tempo di percorrenza è minimo. Anzitutto, sperimentalmente, sappiamo che in ciascuno dei due mezzi la luce si propagerà in linea retta. Dunque cercheremo la curva di tempo minimo nell'insieme  $\mathcal{W}_{AB}^r$  delle spezzate ACB che hanno un vertice C sulla retta  $r$ . Posso identificare ogni elemento di questo insieme con un numero reale  $x$ , l'ascissa di C su  $r$ :  $\mathcal{W} = \{x \in \mathbb{R} : 0 \leq x \leq l\}$  dove  $l$  è la distanza tra le proiezioni ortogonali di A e di B e ho fissato l'origine del sistema di riferimento in O, come in figura.

La quantità che vogliamo rendere minima per questo problema è il tempo di percorrenza della curva, espresso dalla funzione reale di variabile reale  $F : [0, l] \rightarrow \mathbb{R}$  con

$$F(x) = \frac{l_1}{v_1} + \frac{l_2}{v_2} = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (l-x)^2}}{v_2}. \quad (2)$$

Ora ci basta determinare il valore  $\bar{x}$  per cui  $F$  è minima. Ciò richiede anzitutto  $\bar{x}$  estrema, quindi  $F'(\bar{x}) = 0$ , dove

$$F'(x) = \frac{1}{v_1} \frac{x}{l_1} - \frac{1}{v_2} \frac{l-x}{l_2} = \frac{\sin \theta_1}{v_1} - \frac{\sin \theta_2}{v_2}. \quad (3)$$

Abbiamo ottenuto quindi che la richiesta di estimalità per la funzione  $F$  equivale alla richiesta di validità della legge di rifrazione; più chiaramente, il raggio soddisfacente la legge di rifrazione è quello tale per cui il tempo di percorrenza è minimo. Dunque le curve significative per l'ottica geometrica sono caratterizzate da una proprietà di massimo e minimo del funzionale "tempo di percorrenza". Quello che voglio fare, più in generale, è caratterizzare le curve mediante la condizione di estimalità di un assegnato funzionale.

### Fermat e Maupertuis

Consideriamo un mezzo arbitrario purché isotropo, ossia un mezzo le cui proprietà possono variare punto per punto, in cui però la variazione non dipenda dalla direzione.



In ciascun punto  $x$  del mezzo è assegnata una velocità locale di propagazione della luce  $v(x)$ ; in modo equivalente è assegnato un *indice di rifrazione*  $n(x)$

$$n(x) = \frac{c}{v(x)} \tag{4}$$

dove  $c$  indica la velocità della luce nel vuoto. Sia inoltre  $dl$  elemento di linea infinitesimo, ossia tale che  $dl^2 = dx^2 + dy^2 + dz^2$ , dove  $x, y, z$  sono le coordinate cartesiane ortogonali.

L'elemento di linea  $dl$  viene percorso in tempo  $dt$

$$dt = \frac{dl}{v(x)} \quad \text{che equivale a} \quad c dt = n(x) dl \tag{5}$$

Per percorrere una curva  $\gamma$  è necessario un tempo  $T(\gamma)$  dato da

$$T(\gamma) = \int_{\gamma} dt = \int_{\gamma} \frac{dl}{v(x)} \quad \text{ovvero} \quad cT(\gamma) = \int_{\gamma} n(x) dl \tag{6}$$

Si noti che il funzionale  $T(\gamma)$  generalizza il funzionale definito nell'esempio della rifrazione (2).

**PRINCIPIO DI FERMAT.** In un mezzo caratterizzato otticamente da un indice di rifrazione  $n(x)$ , il raggio luminoso  $\bar{\gamma}$  tra due punti  $A$  e  $B$  che si realizza in natura è la curva, tra tutte le curve  $\gamma$  congiungenti i punti, per cui è estrema il tempo di percorrenza, o equivalentemente lo è il funzionale

$$F(\gamma) = \int_{\gamma} n(x) dl \tag{7}$$

detto *cammino ottico*.

Dico estrema di un funzionale una curva  $\bar{\gamma}$  che sia tale da soddisfare  $\delta F(\bar{\gamma}) = 0$ , dove  $\delta$  è l'analogo del differenziale in  $\mathbb{R}^n$ , che ora definiamo. Per farlo ricordiamo che il gradiente  $\nabla f$  di una funzione differenziabile  $f$  è quel vettore per cui vale

$$\langle \nabla f(x), y \rangle = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} f(x + \varepsilon y) . \tag{8}$$

Analogamente quindi definisco  $\delta$ . Dato  $\mathcal{L}[u]$  funzionale del tipo

$$\mathcal{L}[u] = \int_{\Omega} \mathcal{L}(x, u^{(n)}) dx , \tag{9}$$

la *derivata variazionale* di  $\mathcal{L}$  è il vettore  $\delta \mathcal{L}[u] = (\delta_1 \mathcal{L}[u], \dots, \delta_q \mathcal{L}[u])$ , tale per cui

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{L}[f + \varepsilon \eta] = \int_{\Omega} \delta \mathcal{L}[f(x)] \cdot \eta(x) dx \tag{10}$$

dove  $u = f(x)$  è una funzione  $C^\infty$  definita su  $\Omega$ , e  $\eta(x) = (\eta^1(x), \dots, \eta^q(x))$  è una funzione  $C^\infty$  a supporto compatto in  $\Omega$ , in modo tale che  $f + \varepsilon \eta$  soddisfi qualsiasi condizione al bordo imposta allo spazio di funzioni su cui sto estremizzando  $\mathcal{L}$ . In questo caso come prodotto scalare ho utilizzato il prodotto interno di  $L^2$ .

Torniamo ora sui nostri passi. Il principio di Fermat è innovativo, ma il vero passaggio

tra l'ottica geometrica e la meccanica dei corpi si deve a Maupertuis, che in particolare aveva studiato tre analogie:

- il moto rettilineo di un corpo non soggetto a forze, confrontato con la propagazione in linea retta dei raggi luminosi;
- il moto di un corpo che urtando contro una parete rimbalza in modo perfettamente elastico confrontato con la legge di riflessione della luce;
- il moto di un corpo che si muove attraverso mezzi di densità diversa confrontato con la legge di rifrazione della luce.

Quello che egli considerava inaccettabile è che, mentre nei primi due casi l'analogia tra ottica geometrica e meccanica è completa, in quanto luce e corpi materiali seguono sempre il cammino più breve, nell'ultimo la descrizione del moto di un corpo risulta molto più complessa, mentre per la luce si deve modificare il punto di vista, affermando che il raggio luminoso minimizza il tempo, non il percorso. Certamente nei primi due casi la minimizzazione del tempo coincide con quella del percorso, tuttavia perché si dovrebbe preferire il tempo allo spazio?

Qui Maupertuis segue un percorso puramente speculativo, con l'obiettivo dichiarato di individuare una quantità che misuri l'Azione della Natura, e di assoggettare quella quantità a un principio di minimo. Vediamo come è riuscito a trovarla in un caso semplice.

Considero il caso del moto di un punto soggetto a un campo di forze  $F(x)$ , che ammettono energia potenziale  $V(x)$ , perciò il moto soddisfa la legge di Newton

$$m\ddot{x} = -\nabla V \quad (11)$$

Siamo interessati alla traiettoria effettuata dal punto, che risponde alla legge del moto  $x = x(t)$ . Si trova che le traiettorie  $\gamma$  corrispondenti ai moti  $x = x(t)$  soddisfacenti l'equazione di Newton soddisfano un principio variazionale effettivamente analogo al principio di Fermat:

**PRINCIPIO DI MAUPERTUIS-JACOBI.** Le traiettorie di un punto materiale soggetto a un'energia potenziale  $V(x)$ , a una fissata energia  $E$ , sono caratterizzate dalla proprietà di estremalità

$$\delta \int_{\gamma} \sqrt{E - V(x)} dl = 0 \quad (12)$$

In altri termini, la traiettoria di un punto materiale in un potenziale  $V$  a energia  $E$  coincide con il cammino di un raggio luminoso in un mezzo ottico avente indice di rifrazione

$$n(x) = \sqrt{E - V(x)} \quad (13)$$

### Hamilton, Eulero e Lagrange

In questo contesto si inseriscono i risultati ottenuti da Hamilton attorno al 1833, quando giunse alla sua formulazione della meccanica, hamiltoniana appunto. Consideriamo una varietà differenziabile  $X_n$  di coordinate  $x = (x^1, \dots, x^n)$ . Sia poi  $\gamma : x^j = x^j(t)$  una curva su  $X_n$ , dove  $x^j$  sono funzioni di classe  $C^2$  e  $t$  denota un parametro dato. Vestendo ancora per un attimo i panni dei fisici vediamo  $t$  come

il tempo in un sistema dinamico in cui  $(x^1, \dots, x^n)$  sono le coordinate di un punto materiale in moto lungo una traiettoria  $\gamma$  con velocità  $\dot{x} = (\dot{x}^1, \dots, \dot{x}^n)$ . Definisco una funzione  $\mathcal{L}$

$$\mathcal{L} : \mathbb{R}^{2n+1} \rightarrow \mathbb{R}, \quad \mathcal{L}(t, x^j, \dot{x}^j) \tag{14}$$

detta Lagrangiana del sistema. Suppongo  $\mathcal{L}$  di classe  $C^2$ .

PRINCIPIO DI MINIMA AZIONE DI HAMILTON. Per un sistema meccanico conservativo descritto dalla Lagrangiana  $\mathcal{L}(t, x^j, \dot{x}^j)$ , i moti naturali sono le curve  $\gamma : x^j = x^j(t)$  tali per cui il funzionale

$$\mathcal{L}(\gamma) = \int_{\gamma} \mathcal{L}(t, x^j, \dot{x}^j) dt, \tag{15}$$

detto funzionale d'azione, raggiunge valori estremi.

In generale, dati  $P_1, P_2$  punti della curva  $x^j(t)$  per valori di  $t$  rispettivamente  $t_1, t_2$ , l'integrale d'azione lungo  $\gamma$  sarà descritto da:

$$\mathcal{L}(\gamma) = \int_{t_1}^{t_2} \mathcal{L}(t, x^j, \dot{x}^j) dt, \tag{16}$$

quindi  $\mathcal{L}(\gamma)$  dipende dalla scelta della curva  $\gamma$  che unisce i punti  $P_1, P_2$ .

Ora, i moti in natura avvengono lungo traiettorie che rendono stazionario il funzionale d'azione, e allora quali sono le caratteristiche di queste curve fortunate? Quello che si ottiene con i procedimenti seguenti è un risultato di grande importanza nella teoria dei principi variazionali. Di seguito l'enunciato.

PROPOSIZIONE 1. *Affinché la curva  $\gamma$ , definita in  $X_n$  come sopra, sia estrema per l'integrale d'azione (16) è necessario che le funzioni  $x^j(t)$  soddisfino un sistema di  $n$  equazioni differenziali del secondo ordine, dette equazioni di Eulero-Lagrange.*

Sia  $\gamma : x^j = x^j(t)$  curva di classe  $C^1$  in  $X_n$ , tale che unisca i punti  $P_1, P_2$ , che raggiunge quando  $t$  assume i valori  $t_1, t_2$ . Costruisco una famiglia di curve a un parametro

$$\gamma(k) : x^j(t, k) = x^j(t) + k\zeta^j(t) \tag{17}$$

dove le  $\zeta^j$  sono curve  $C^1$  in  $t$  che rispettino la condizione  $\zeta^j(t_1) = \zeta^j(t_2) = 0$ , in questo modo per ogni valore di  $k$ ,  $\gamma(k)$  passa per i punti  $P_1, P_2$ . Se le  $\zeta^j$  sono tali da rispettare la condizione al bordo, allora possono essere scelte in modo arbitrario. Noto inoltre che la curva iniziale appartiene in modo banale alla famiglia per il valore del parametro  $k = 0$ .

Suppongo che la curva estrema per il problema sia  $\gamma = \gamma(0)$  e mi chiedo quali siano le condizioni che la rendono appunto estrema. Per come ho definito la famiglia di curve, le  $\zeta^j$  rappresentano le variazioni, modulate da  $k$ , di  $x^j$  viste come funzioni che definiscono  $\gamma$ . Differenzio rispetto al tempo la relazione (17) che definisce la curva per ottenere una relazione tra le derivate:

$$\dot{x}^j(t, k) = \dot{x}^j(t) + k\dot{\zeta}^j(t) \tag{18}$$

L'integrale d'azione valutato lungo la curva  $C(k)$  sarà di conseguenza:

$$\mathcal{L}(k) = \mathcal{L}(\gamma(k)) = \int_{t_1}^{t_2} \mathcal{L} \{t, x^j + k\zeta^j(t), \dot{x}^j(t) + k\dot{\zeta}^j(t)\} dt \quad (19)$$

Esprimo  $\mathcal{L}(k)$  con il suo sviluppo di Taylor, arrestato al secondo ordine e centrato in  $k = 0$

$$\mathcal{L}(k) = \mathcal{L}(0) + k \left( \frac{d\mathcal{L}}{dk} \right) \Big|_{k=0} + \frac{1}{2} k^2 \left( \frac{d^2\mathcal{L}}{dk^2} \right) \Big|_{k=0} + o(k^2) \quad (20)$$

Supposta  $C$  estrema,  $\mathcal{L}(k)$  ha un estremo in  $k = 0$ , quindi

$$\delta\mathcal{L} = \left( \frac{d\mathcal{L}}{dk} \right) \Big|_{k=0} = 0 \quad (21)$$

deve essere soddisfatta. Voglio tradurre questa condizione in un criterio sulle funzioni che definiscono la curva, perciò calcolo  $\delta\mathcal{L}$ .

$$\left( \frac{d\mathcal{L}}{dk} \right) \Big|_{k=0} = \int_{t_1}^{t_2} \left\{ \frac{\partial\mathcal{L}}{\partial x^j} \zeta^j + \frac{\partial\mathcal{L}}{\partial \dot{x}^j} \dot{\zeta}^j \right\} dt \quad (22)$$

Il primo termine dell'integranda è risolvibile per parti nel modo seguente

$$\int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial x^j} \zeta^j dt = \left[ \zeta^j \int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial x^j} dt \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \left\{ \dot{\zeta}^j \int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial x^j} dt \right\} dt = - \int_{t_1}^{t_2} \left\{ \dot{\zeta}^j \int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial x^j} dt \right\} dt \quad (23)$$

in cui la seconda uguaglianza vale per le condizioni al bordo imposte sulle variabili  $\zeta^j$ . Perciò, risostituendo in (22) e imponendo  $\delta\mathcal{L} = 0$  so che deve valere

$$\int_{t_1}^{t_2} \dot{\zeta}^j \left\{ \frac{\partial\mathcal{L}}{\partial \dot{x}^j} - \int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial x^j} dt \right\} dt = 0, \quad (24)$$

che possiamo riscrivere introducendo, per brevità, nuove variabili  $\Phi_j$  in questo modo:

$$\int_{t_1}^{t_2} \dot{\zeta}^j \Phi_j dt = 0 \quad \text{dove} \quad \Phi_j = \frac{\partial\mathcal{L}}{\partial \dot{x}^j} - \int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial x^j} dt. \quad (25)$$

Siano poi  $c_j$  per  $j \in \{1, \dots, n\}$  costanti definite dalle relazioni equivalenti:

$$c_j(t_2 - t_1) = \int_{t_1}^{t_2} \Phi_j dt \quad \int_{t_1}^{t_2} \{\Phi_j - c_j\} dt = 0. \quad (26)$$

Ora, so che la (25) deve valere per ogni scelta delle  $\zeta^j(t)$ , su cui ho posto l'unica condizione che si annullino ai bordi di  $[t_1, t_2]$ , perciò posso sceglierle in modo arbitrario purché rispettino le condizioni al bordo, e cercare così una condizione sulle  $\Phi_j$ . Quindi definisco le funzioni che determinano la famiglia di curve  $\forall j$  come

$$\zeta^j(t) = \int_{t_1}^t \{\Phi_j - c_j\} dt \quad \text{perciò} \quad \dot{\zeta}^j(t) = \Phi_j - c_j. \quad (27)$$

Abbiamo già verificato che  $\zeta^j(t_1) = \zeta^j(t_2) = 0$ . In conclusione, la condizione (21) risulta quindi equivalente a

$$\int_{t_1}^{t_2} \sum_{j=1}^n \Phi_j(\Phi_j - c_j) dt = \int_{t_1}^{t_2} \sum_{j=1}^n (\Phi_j - c_j)(\Phi_j - c_j) dt = 0 \quad (28)$$

Ora l'argomento dell'integrale è dato da una somma di quadrati, perciò si annulla solo se  $\Phi_j - c_j = 0$  ossia

$$\frac{\partial \mathcal{L}}{\partial \dot{x}^j} - \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial x^j} dt = c_j \quad (29)$$

che differenziata rispetto al tempo diventa

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}^j} - \frac{\partial \mathcal{L}}{\partial x^j} = 0 \quad (30)$$

Abbiamo perciò concluso che affinché la curva  $\gamma$  sia tale da rendere stazionario l'integrale d'azione  $\mathcal{L}$  è necessario che le funzioni  $x^j(t)$  che definiscono la curva soddisfino le equazioni (30), che si dicono *equazioni di Eulero-Lagrange*. Si tratta di un sistema di  $n$  equazioni differenziali ordinarie del secondo ordine fondamentali per la descrizione di modelli fisici, in quanto per un sistema conservativo sono equivalenti al secondo principio della dinamica.

### Bibliografia

- [LDR] LOVELOCK, DAVID, AND HANNO RUND (1989), *Tensors, differential forms, and variational principles*, Courier Corporation, .



# STRANEZZE $p$ -ADICHE

SIMONE PULEIO

I numeri  $p$ -adici, con  $p$  numero primo, vennero introdotti per la prima volta alla fine dell'800' ad opera del matematico tedesco Kurt Hensel.

Storicamente, il punto di partenza di Hensel furono certe analogie osservate tra l'anello degli interi  $\mathbb{Z}$  e il suo campo delle frazioni  $\mathbb{Q}$ , e tra l'anello  $\mathbb{C}[X]$  dei polinomi a coefficienti in  $\mathbb{C}$  e il suo campo delle frazioni  $\mathbb{C}(X)$ , che suggerivano di poter generalizzare l'idea delle serie di Taylor al caso di numeri interi e dei numeri razionali.

Esistono vari modi di procedere nella costruzione dell'insieme  $\mathbb{Q}_p$  dei numeri  $p$ -adici: è possibile, per esempio, costruire  $\mathbb{Q}_p$  in analogia con l'insieme dei numeri reali  $\mathbb{R}$ . Con questo approccio, si parte dall'osservare che la costruzione di  $\mathbb{R}$  come completamento di  $\mathbb{Q}$ , rispetto alla distanza euclidea, è una costruzione arbitraria, cioè che sarebbe possibile in linea di principio considerare un'altra metrica su  $\mathbb{Q}$ , e poi completare di conseguenza lo spazio.

Dopo la costruzione, si riprenderanno alcuni risultati classici sulla convergenza di successioni e serie a valori reali, e si interpreteranno nello spazio  $\mathbb{Q}_p$ , in analogia con il caso reale.

## Costruzione di $\mathbb{Q}_p$

DEFINIZIONE 1. Dato un campo  $\mathbb{K}$ , un valore assoluto su  $\mathbb{K}$  è una funzione  $|\cdot| : \mathbb{K} \rightarrow \mathbb{R}_{\geq 0}$  tale che, per ogni  $x, y \in \mathbb{K}$

- (1)  $|x| = 0$  se e solo se  $x = 0$
- (2)  $|x \cdot y| = |x| \cdot |y|$
- (3)  $|x + y| \leq |x| + |y|$

Inoltre, se per ogni  $x, y \in \mathbb{K}$  vale  $|x + y| \leq \max\{|x|, |y|\}$  il valore assoluto è detto non-Archimedeo.

Per esempio, per  $\mathbb{K} = \mathbb{R}$ , due possibili scelte sono il valore assoluto standard euclideo  $|\cdot|_e$  o il valore assoluto banale  $|\cdot|_b$  definiti come:

$$|x|_e = \begin{cases} x & \text{se } x \geq 0 \\ -x & \text{se } x < 0 \end{cases} \quad |x|_b = \begin{cases} 1 & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

Ci restringiamo ora al caso  $\mathbb{K} = \mathbb{Q}$ . Vogliamo definire un valore assoluto  $|\cdot|$  su  $\mathbb{Q}$ , e quindi una nuova metrica, non equivalente al valore assoluto standard, cioè tale che

non esista una costante  $k \in \mathbb{R}_{\geq 0}$  per cui  $|\cdot|^k = |\cdot|_e$ .

Sia d'ora in poi  $p$  un numero primo fissato.

DEFINIZIONE 2. La valutazione  $p$ -adica su  $\mathbb{N}$  è la funzione  $v_p : \mathbb{N} \setminus \{0\} \rightarrow \mathbb{R}$  tale che

$$v_p(n) = \max\{k \in \mathbb{N} : p^k | n\}$$

Si estende a  $v_p : \mathbb{Z} \rightarrow \mathbb{R}$  ponendo  $v_p(-n) := v_p(n)$  per ogni  $n \in \mathbb{N}$ .

Si estende a  $v_p : \mathbb{Q} \rightarrow \mathbb{R}$  ponendo  $v_p\left(\frac{a}{b}\right) := v_p(a) - v_p(b)$ .

Si pone per convenzione  $v_p(0) = +\infty$ , dal momento che  $p^k | 0$  per ogni  $k \in \mathbb{N}$ .

DEFINIZIONE 3. Il valore assoluto  $p$ -adico è la funzione  $|\cdot|_p : \mathbb{Q} \rightarrow \mathbb{R}$  tale che

$$|x|_p = \begin{cases} p^{-v_p(x)} & \text{se } x \in \mathbb{Q} \setminus \{0\} \\ 0 & \text{se } x = 0 \end{cases}$$

PROPOSIZIONE 4.  $|\cdot|_p$  è valore assoluto non-Archimedeo su  $\mathbb{Q}$ .

*Dimostrazione.* (Idea) Si sfruttano le proprietà della valutazione  $p$ -adica su  $\mathbb{Q}$  :

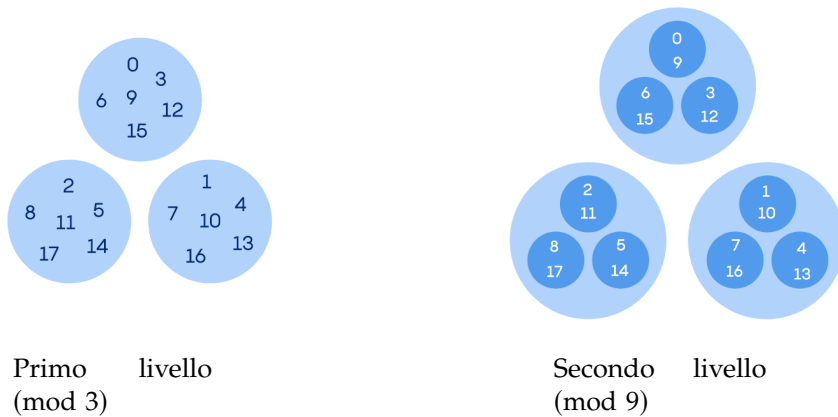
- $v_p(xy) = v_p(x) + v_p(y)$
- $v_p(x + y) \geq \min\{v_p(x), v_p(y)\}$

□

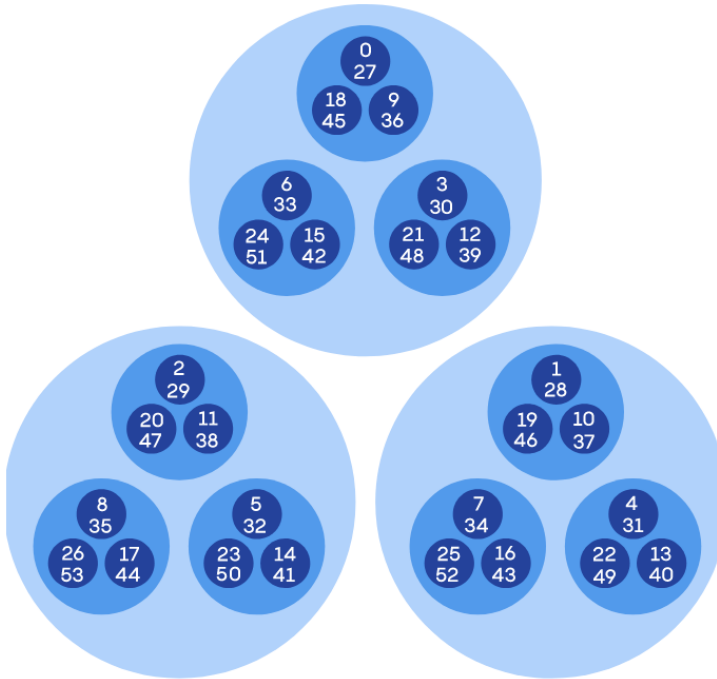
Il valore assoluto  $|\cdot|_p$  su  $\mathbb{Q}$  induce una distanza  $d_p(x, y) = |x - y|_p$  su  $\mathbb{Q}$ , detta metrica  $p$ -adica, e quindi il nuovo spazio metrico  $(\mathbb{Q}, d_p)$ .

Intuitivamente, si può pensare che il valore assoluto  $|\cdot|_p$  sia una "misura" della divisibilità per  $p$ , nel senso che tanto più un numero  $n$  è divisibile per  $p$ , tanto più  $|n|_p$  sarà piccolo.

Analogamente, se ci si restringe al caso di  $\mathbb{Z}$ , si può pensare che con la metrica  $d_p$ , gli interi siano disposti su circonferenze concentriche su più livelli come in figura, in cui interi sullo stesso livello hanno la stessa distanza. Tanto più si procede nei livelli, quanto più gli interi sono congrui modulo una potenza "grande" di  $p$ . Per esempio per  $p = 3$ ,







Terzo livello (mod 27)

PROPOSIZIONE 5.  $(\mathbb{Q}, d_p)$  è uno spazio metrico non completo.

TEOREMA 6. Sia  $\mathbb{K}$  campo con valore assoluto  $|\cdot|$ , allora esiste  $\mathbb{K}'$  campo completo con valore assoluto  $|\cdot|'$  che estende  $\mathbb{K}$ .

Inoltre, il completamento  $\mathbb{K}'$  è unico a meno di isomorfismo, e la restrizione di  $|\cdot|'$  a  $\mathbb{K}$  coincide con  $|\cdot|$

*Dimostrazione.* (Idea) Analoga alla costruzione di  $\mathbb{R}$  tramite successioni di Cauchy, con alcuni accorgimenti algebrici per mostrare che lo spazio completo ottenuto è un campo.  $\square$

In analogia al caso standard, cioè a partire da  $\mathbb{Q}$  con la metrica euclidea, si ottiene il campo dei reali  $\mathbb{R}$ , completo rispetto alla distanza euclidea. Invece, se si considera  $\mathbb{Q}$  con la metrica indotta dal valore assoluto  $p$ -adico  $|\cdot|_p$ , per il teorema precedente si ottiene un nuovo campo, detto campo dei numeri  $p$ -adici  $\mathbb{Q}_p$ , completo rispetto all'estensione della metrica  $d_p$ .

In realtà, nel 1916 *Ostrowski* provò un risultato che afferma che  $\mathbb{R}$  e  $\mathbb{Q}_p$ , al variare di  $p$  primo, sono gli unici completamenti possibili di  $\mathbb{Q}$ .

TEOREMA. (*Ostrowski*) Ogni valore assoluto non banale su  $\mathbb{Q}$  è equivalente al valore assoluto standard euclideo o ad un valore assoluto  $p$ -adico per qualche  $p$  primo.

### Stranezze

Richiamiamo prima alcune definizioni utili.

Sia  $(X, d)$  spazio metrico :

- Una successione  $\{a_n\}_{n \in \mathbb{N}}$  in  $X$  è di Cauchy se, per ogni  $\varepsilon > 0$  esiste  $N \in \mathbb{N}$  tale che per ogni  $n, m \geq N$  si abbia  $d(a_n, a_m) < \varepsilon$ .
- $X$  è detto completo se ogni successione di Cauchy è convergente.

STRANEZZA 1. Fissato  $p$  primo, la successione  $\{p^n\}_{n \in \mathbb{N}}$  converge a 0 in  $\mathbb{Q}_p$

*Dimostrazione.* Fissato  $\varepsilon > 0$ , è sufficiente mostrare che esiste  $N \in \mathbb{N}$  tale che  $d_p(p^n, 0) = |p^n|_p < \varepsilon$  per ogni  $n \geq N$ .

Sia  $N \in \mathbb{N}$  tale che  $\frac{1}{N} < \varepsilon$ , quindi

$$\frac{1}{p^n} < \frac{1}{p^N} < \frac{1}{N} < \varepsilon \text{ per ogni } n \geq N$$

Da cui la tesi poiché  $|p^n|_p = p^{-n} < \varepsilon$ . □

STRANEZZA 2. Sia  $\{a_n\}_{n \in \mathbb{N}}$  successione in  $\mathbb{Q}_p$ , allora  $\{a_n\}_{n \in \mathbb{N}}$  successione di Cauchy se e solo se

Per ogni  $\varepsilon > 0$ , esiste  $N \in \mathbb{N}$  tale che  $|a_{n+1} - a_n|_p < \varepsilon$  per ogni  $n \geq N$

*Dimostrazione.*  $\Rightarrow$ ) Vera in qualsiasi spazio metrico completo; infatti ogni  $\{a_n\}_{n \in \mathbb{N}}$  successione di Cauchy è convergente, quindi la successione delle differenze  $\{a_{n+1} - a_n\}_{n \in \mathbb{N}}$  converge a 0, da cui la tesi.

$\Leftarrow$ ) Fissato  $\varepsilon > 0$ , è sufficiente mostrare che esiste  $N \in \mathbb{N}$  tale che  $\forall n, m \geq N$  si ha  $|a_n - a_m|_p < \varepsilon$ .

Sia  $N \in \mathbb{N}$ , per ipotesi, tale che per ogni  $n \geq N$  si abbia  $|a_{n+1} - a_n|_p < \varepsilon$ , e sia  $m \geq N$ . Infine suppongo  $n > m$ , cioè  $n = m + k$  per qualche  $k \in \mathbb{N}$ , allora

$$|a_n - a_m|_p = |(a_{m+k} - a_{m+k-1}) + \dots + (a_{m+1} - a_m)|_p$$

A questo punto, in una dimostrazione standard di Analisi 1, utilizzeremmo la disuguaglianza triangolare per maggiorare il termine destro dell'uguaglianza, ma in questo caso è uno strumento troppo debole per poter concludere. Invece per la proprietà non-Archimedeica si ottiene :

$$\begin{aligned} |a_n - a_m|_p &= |(a_{m+k} - a_{m+k-1}) + \dots + (a_{m+1} - a_m)|_p \\ &\leq \max\{|a_{m+k} - a_{m+k-1}|_p, \dots, |a_{m+1} - a_m|_p\} \\ &= |a_{m+i_0} - a_{m+i_0-1}|_p < \varepsilon \text{ per qualche } i_0 \in \{1, \dots, k\}. \end{aligned}$$

□

STRANEZZA 3. Sia  $\{a_n\}_{n \in \mathbb{N}}$  successione in  $\mathbb{Q}_p$ , allora  $\sum_{n=0}^{+\infty} a_n$  converge se e solo se  $\{a_n\}_{n \in \mathbb{N}}$  converge a 0 in  $\mathbb{Q}_p$ .

*Dimostrazione.*  $\Rightarrow$ ) Vera in qualsiasi spazio metrico.

$\Leftarrow$ ) Per ipotesi  $\{a_n\} \rightarrow 0$  in  $\mathbb{Q}_p$ , cioè fissato  $\varepsilon > 0$ , esiste  $N \in \mathbb{N}$  tale che  $|a_n|_p < \varepsilon$  per

ogni  $n \geq N$ ; in particolare anche  $|a_{n+1}|_p < \varepsilon$ .

Indicando con  $S_n = \sum_{i=0}^n a_i$  l' $n$ -esima somma parziale, si ha

$$|S_{n+1} - S_n|_p = |a_{n+1}|_p < \varepsilon \text{ per ogni } n \geq N.$$

Per la stranezza (2), la successione delle somme parziali  $\{S_n\}_{n \in \mathbb{N}}$  è di Cauchy in  $\mathbb{Q}_p$  completo, quindi  $\sum_{n=0}^{+\infty} a_n$  converge in  $\mathbb{Q}_p$ .  $\square$

Quindi, a differenza delle serie ambientate nel campo reale, abbiamo ottenuto una condizione necessaria e sufficiente per determinare la convergenza di una serie.

Inoltre, per immediata applicazione del risultato precedente, si ha che la serie  $\sum_{n=0}^{+\infty} p^n$  converge in  $\mathbb{Q}_p$ , poiché per la stranezza (1) il termine generale  $p^n \rightarrow 0$  per  $n \rightarrow +\infty$ .

In particolare, come al caso reale, si ha  $\sum_{n=0}^{+\infty} p^n = \frac{1}{1-p}$ .

Infatti l' $n$ -esima somma parziale è pari a

$$S_n = \sum_{i=0}^n p^i = \frac{1 - p^{n+1}}{1 - p}$$

Allora,

$$\lim_{n \rightarrow +\infty} S_n = \lim_{n \rightarrow +\infty} \frac{1 - p^{n+1}}{1 - p} = \frac{1}{1 - p} \cdot \lim_{n \rightarrow +\infty} (1 - p^{n+1}) = \frac{1}{1 - p}.$$

Più in generale, converge in  $\mathbb{Q}_p$  ogni serie della forma

$$\sum_{n=n_0}^{+\infty} a_n p^n \text{ con } n_0 \in \mathbb{Z}, a_n \in \{0, \dots, p-1\}$$

Le serie della forma  $\sum_{n=n_0}^{+\infty} a_n p^n$  con  $n_0 \in \mathbb{Z}$ ,  $a_n \in \{0, \dots, p-1\}$  sono dette espansioni  $p$ -adiche, e costituiscono di fatto una rappresentazione standard per i numeri  $p$ -adici.

### Bibliografia

- [P] POMERANTZ, A. (2020), *An Introduction to the  $p$ -adic numbers*  
 [G] Q. GOUVÊA (2003),  *$p$ -adic Numbers, An Introduction*



# UNA PUNTA DI IPOTESI DEL CONTINUO

PAOLO BOLDRINI

L'8 agosto 1900 si tiene a Parigi il congresso internazionale dei matematici. Al convegno prende parte uno dei matematici di punta del periodo: David Hilbert. Hilbert tiene un discorso solenne nel quale propone 10 problemi da lui ritenuti i più significativi ancora irrisolti (in seguito come ben sapete la lista è stata estesa a 23). Tra questi il primo enunciato è la celebre ipotesi del continuo.

Come noto a tutti gli studenti di matematica, l'insieme  $\mathbb{N}$  dei numeri naturali è molto più "piccolo" dell'insieme  $\mathbb{R}$  dei numeri reali. Più precisamente la cardinalità di  $\mathbb{N}$  è strettamente minore della cardinalità di  $\mathbb{R}$ . Questo fatto fu scoperto da Georg Cantor. Proprio Cantor fu il primo a formulare la seguente domanda: *c'è qualcosa in mezzo? Esiste un insieme che abbia cardinalità strettamente compresa tra quella di  $\mathbb{N}$  e quella di  $\mathbb{R}$ ? La (eventuale) risposta negativa a tale domanda è chiamata "ipotesi del continuo".*

Illustri matematici cercarono invano di trovare una risposta, positiva o negativa che fosse. A posteriori il loro fallimento non è da biasimare: oggi sappiamo che una risposta non c'è!

Nel 1940 Gödel dimostrò che non è possibile dimostrare la falsità dell'ipotesi del continuo. Nel 1963 Cohen dimostrò invece che non è possibile dimostrare la sua verità<sup>1</sup>.

Ma quindi, se non possiamo dare una risposta, di cosa parliamo nelle prossime pagine? L'ipotesi del continuo come l'abbiamo enunciata è estremamente generale, parla di insiemi privi di struttura. Se però restringiamo il campo di indagine a insiemi dotati di una topologia particolarmente maneggevole possiamo fornire una risposta ben più soddisfacente.

La sezione 1 riporta alcune proprietà elementari dei numeri cardinali. La sezione 2 invece fornisce le nozioni di topologia che ci saranno utili. Per i più curiosi i risultati citati possono essere trovati in (rispettivamente) "Martin Hils e Loeser François. *A first journey through logic*. American mathematical society, 2019" e in "Marco Manetti. *Topologia*. Springer, 2014". Nelle sezioni 3 e 4 ci addentriamo nel cuore del problema, il materiale esposto qui è in gran parte tratto da "Alexander S. Kechris. *Classical descriptive set theory*. Springer, 1995".

---

<sup>1</sup>Più precisamente l'ipotesi del continuo e la sua negazione sono entrambe compatibili con il sistema di assiomi ZFC, ovvero il sistema di assiomi generalmente utilizzato per fare matematica. Per una discussione più dettagliata si consulti "Kenneth Kunen. *Set theory: An introduction to independence proofs*. Elsevier, 2006"

## 1. Numeri cardinali

Intuitivamente un numero cardinale può essere visto come un insieme "prototipo per la cardinalità". Un po' impropriamente possiamo pensare un cardinale come un rappresentante canonico di una classe di equivalenza in *Set* modulo la relazione "essere in biezione".

Non è data una definizione formale di numero cardinale poiché porterebbe troppo lontano. Sono però elencati alcuni fatti che ci saranno utili in seguito.

- (1) Ogni insieme è in biezione con esattamente un cardinale (per garantire questo fatto è necessario l'assioma della scelta). Dato un insieme  $X$  scriviamo  $|X|$  per indicare l'unico cardinale con esso in biezione.
- (2) Possiamo eseguire operazioni sui numeri cardinali. In particolare sarà rilevante l'operazione  $\kappa^\lambda$ , dove  $\kappa$  e  $\lambda$  sono cardinali.
- (3) Fissato un cardinale  $\kappa$ , se  $\lambda \leq \mu$  si ha che  $\lambda^\kappa \leq \mu^\kappa$ .
- (4)  $(\kappa^\lambda)^\mu = \kappa^{\lambda\mu}$ .
- (5) Per ogni cardinale  $\kappa$ ,  $\kappa < 2^\kappa$ .
- (6) Dato un insieme  $X$ ,  $|2^X| = 2^{|X|}$ .
- (7) Siano  $X$  e  $Y$  insiemi, se esiste un'iniezione da  $X$  in  $Y$ ,  $|X| \leq |Y|$ , viceversa se esiste una suriezione  $|X| \geq |Y|$ .
- (8) Se  $|X| \leq |Y|$  e  $|Y| \leq |X|$  allora  $|X| = |Y|$ . Questo risultato è noto come teorema di Cantor-Bernstein.
- (9) Dato un cardinale  $\kappa$  esiste sempre il minimo cardinale strettamente maggiore di esso. Denotiamo  $\kappa^+$  tale cardinale, detto successore di  $\kappa$ .
- (10) Poniamo  $\aleph_0 = |\mathbb{N}|$  e  $\aleph_1 = \aleph_0^+$ .

LEMMA 1. (1) Per ogni insieme  $X$ ,  $|\mathcal{P}(X)| = 2^{|X|}$ .

$$(2) \aleph_0^{\aleph_0} = 2^{\aleph_0}.$$

$$(3) |\mathbb{R}| = 2^{\aleph_0}.$$

*Dimostrazione.* (1) La funzione che ad ogni sottoinsieme di  $X$  associa la sua funzione caratteristica è una biezione.

$$(2) \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0^2} = 2^{\aleph_0} \leq \aleph_0^{\aleph_0}.$$

(3) Per mostrare che  $2^{\aleph_0} \leq |\mathbb{R}|$ , è sufficiente osservare che possiamo associare in modo iniettivo ogni sequenza  $(x_0, x_1, \dots)$  al numero reale  $0.x_0x_1\dots$ . Viceversa, poiché un numero reale non è altro che una successione di numeri razionali, si ha che  $|\mathbb{R}| \leq \aleph_0^{\aleph_0} = 2^{\aleph_0}$ .  $\square$

Con il linguaggio appena introdotto l'ipotesi del continuo può essere riformulata nel seguente modo:

$$2^{\aleph_0} = \aleph_1??$$

## 2. Richiami di topologia

DEFINIZIONE 2. Uno spazio topologico è detto separabile se contiene un sottoinsieme numerabile denso (i.e. tale da intersecare ogni aperto non vuoto).

DEFINIZIONE 3. Uno spazio topologico è detto di Hausdorff se punti distinti possono essere separati da aperti disgiunti.

DEFINIZIONE 4. Uno spazio topologico è detto  $T_4$  se è di Hausdorff e per ogni chiuso  $C$  e per ogni aperto  $U \supseteq C$  esiste un aperto  $V$  tale che  $C \subseteq V \subseteq \bar{V} \subseteq U$ .

DEFINIZIONE 5. Uno spazio topologico è detto compatto se da ogni ricoprimento aperto dello spazio è possibile estrarre un sottoricoprimento finito.

DEFINIZIONE 6. Uno spazio topologico è detto perfetto se ogni suo punto è di accumulazione.

DEFINIZIONE 7. Un punto  $x \in (X, \tau)$  è detto di condensazione se ogni suo intorno è non numerabile (i.e. ha cardinalità  $> \aleph_0$ ).

DEFINIZIONE 8. Uno spazio metrico è detto completo se ogni sua successione di Cauchy è convergente.

TEOREMA 9. Ogni spazio metrico è  $T_4$ , in particolare è di Hausdorff.

TEOREMA 10. Una funzione continua da un compatto in un Hausdorff è chiusa, in particolare se è biettiva è un omeomorfismo.

TEOREMA 11. Ogni spazio metrico separabile ammette base numerabile.

### 3. Spazi polacchi

DEFINIZIONE 12. Uno spazio topologico  $(X, \tau)$  è detto polacco se è separabile e completamente metrizzabile (i.e. ammette una metrica completa compatibile con la topologia).

Vediamo qualche esempio per prendere familiarità con il concetto.

- (1) Ogni insieme finito o numerabile con la topologia discreta è polacco.
- (2) Ogni spazio di Banach separabile, con la topologia indotta dalla norma, è polacco (in effetti Banach stesso era polacco).
- (3) Ogni chiuso in uno spazio polacco è polacco (la separabilità è ereditaria negli spazi metrici e i chiusi sono completi)<sup>2</sup>.

DEFINIZIONE 13. L'insieme di Cantor  $\mathcal{C} \subseteq \mathbb{R}$  è definito ricorsivamente come segue:

$$C_0 = [0, 1].$$

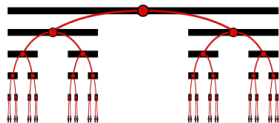
$$C_1 = [0, 1/3] \cup [2/3, 1].$$

...

$$\mathcal{C} = \bigcap_{n \in \mathbb{N}} C_n.$$

PROPOSIZIONE 14.  $\mathcal{C}$  è l'insieme degli elementi compresi tra 0 e 1 che possono essere scritti con uno sviluppo ternario privo di 1, ovvero i numeri della forma  $\sum_{n=0}^{\infty} x_n 3^{-n-1}$  con  $x_n \in \{0, 2\}$ .

<sup>2</sup>Più in generale si può dimostrare che un sottoinsieme di uno spazio polacco è polacco se e solo se è  $G_\delta$ , ovviamente nel caso di insiemi non chiusi la metrica non potrà essere quella indotta.



*Dimostrazione.* (Idea) Al passo 1 vengono eliminati i numeri che hanno 1 come prima cifra ternaria (si noti che  $1/3$  non è eliminato poiché ammette la scrittura  $0.\overline{2}$ ). In generale al passo  $n$ -simo sono eliminati i numeri che hanno 1 come  $n$ -sima cifra ternaria.  $\square$

**PROPOSIZIONE 15.**  $\mathcal{C}$  è omeomorfo a  $2^{\mathbb{N}}$  con la topologia prodotto<sup>3</sup> (2 è da intendersi dotato della topologia discreta).

*Dimostrazione.* La mappa da  $2^{\mathbb{N}}$  a  $\mathcal{C}$

$$(x_0, x_1, \dots) \longrightarrow \sum_{n=0}^{\infty} 2x_n 3^{-n-1}$$

è biettiva in virtù della precedente proposizione. Inoltre è continua. Infatti sia  $(x_0, \dots) \in \mathcal{C}$  e sia  $\varepsilon > 0$ . Si consideri la bolla centrata in  $\sum_{n=0}^{\infty} 2x_n 3^{-n-1}$  di raggio  $\varepsilon$ . Per la convergenza della serie esiste  $N$  tale che  $2 \sum_{n=N}^{\infty} 3^{-n-1} < \varepsilon$ , dunque se  $y \in \prod_{n \in \mathbb{N}} U_n$  con  $U_n = \{x_n\}$  per  $n \leq N$  e  $U_n = 2$  per  $n > N$  la sua immagine è nella bolla fissata. Infine essendo una mappa da un compatto in uno spazio di Hausdorff è chiusa, dunque un omeomorfismo.  $\square$

D'ora in poi identificheremo sistematicamente  $\mathcal{C}$  e  $2^{\mathbb{N}}$ .

Osserviamo che  $\mathcal{C}$  è perfetto. Infatti dato  $x = (x_0, \dots) \in \mathcal{C}$ , un suo intorno elementare è della forma  $\prod_{n \in \mathbb{N}} U_n$ , con  $U_n$  definitivamente (diciamo per  $n > N$ ) uguale a 2. Ma allora  $y = (x_0, \dots, x_N, y_{N+1}, \dots)$  è un elemento dell'intorno ed è distinto da  $x$  a patto di prendere  $y_{N+1} \neq x_{N+1}$ .

#### 4. L'ipotesi del continuo in spazi polacchi

**TEOREMA 16.** Sia  $X$  uno spazio polacco, se  $|X| > \aleph_0$  allora  $|X| = 2^{\aleph_0}$

Dividiamo la dimostrazione del teorema in tre passaggi:

- (1) Gli spazi polacchi hanno cardinalità al più  $2^{\aleph_0}$ .
- (2) Ogni spazio polacco perfetto non vuoto contiene una copia omeomorfa di  $\mathcal{C}$ .
- (3) Ogni spazio polacco può essere decomposto in un insieme perfetto e un insieme numerabile.

**TEOREMA 17.** Ogni spazio polacco ha cardinalità al più  $2^{\aleph_0}$ .

*Dimostrazione.* Sia  $X$  uno spazio polacco con metrica  $d$  completa e  $S$  un suo sottoinsieme denso e numerabile. Si consideri  $C(S)$ , l'insieme delle successioni di Cauchy

<sup>3</sup>La topologia prodotto su  $\prod_{n \in \mathbb{N}} X_n$  è la topologia più debole che rende continue tutte le proiezioni. Può essere caratterizzata tramite la base  $\{\prod_{n \in \mathbb{N}} U_n\}$  con  $U_n$  aperto in  $X_n$  e definitivamente uguale a  $X_n$ .



a valori in  $S$ . Ad ogni successione di Cauchy possiamo associare il punto di  $X$  a cui converge. Dal momento che  $S$  è denso in  $X$  tale mappa è suriettiva, dunque  $|X| \leq |C(S)| \leq |S^{\mathbb{N}}| = \aleph_0^{\aleph_0} = 2^{\aleph_0}$ , dove nell'ultima uguaglianza si è usato il lemma 1.2.  $\square$

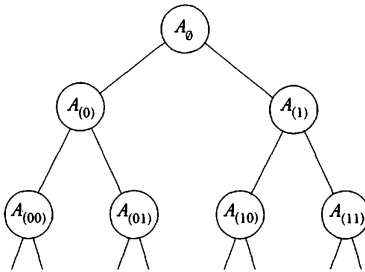
DEFINIZIONE 18. Denotiamo con  $2^n$  l'insieme delle  $n$ -uple di 0 e 1. Poniamo  $2^{<\mathbb{N}} = \bigcup_{n \in \mathbb{N}} 2^n$ . Infine usiamo la notazione  $2^{\mathbb{N}}$  per indicare le successioni a valori in  $\{0, 1\}$ . Per ogni  $s \in 2^{<\mathbb{N}}$ ,  $l(s)$  è il numero naturale tale che  $s \in 2^{l(s)}$ .

Dati due elementi  $s$  e  $t$  in  $2^{<\mathbb{N}}$ ,  $s \star t$  è l'elemento ottenuto congiungendo le due sequenze nell'ordine. Nel caso  $t$  sia una sequenza di un solo valore, ad esempio  $t = (0)$ , si scrive più semplicemente  $s \star 0$ .

Si noti che in particolare  $2^0$  è costituito dalla sola successione vuota:  $2^0 = \{\emptyset\}$ .

DEFINIZIONE 19. Uno schema di Cantor in un insieme  $X$  è una famiglia  $(A_s)_{s \in 2^{<\mathbb{N}}}$  di sottoinsiemi di  $X$  tale che:

- (1)  $A_{s \star 0} \cap A_{s \star 1} = \emptyset$  per  $s \in 2^{<\mathbb{N}}$ .
- (2)  $A_{s \star i} \subseteq A_s$  per  $s \in 2^{<\mathbb{N}}$  e  $i \in \{0, 1\}$ .



Si noti che la costruzione di  $\mathcal{C}$  ricalca uno schema di Cantor in  $\mathbb{R}$ .

TEOREMA 20. Sia  $X$  uno spazio polacco perfetto non vuoto. Allora esiste una copia omeomorfa di  $\mathcal{C}$  dentro  $X$ . In particolare ogni spazio polacco perfetto non vuoto ha cardinalità almeno  $2^{\aleph_0}$ .

*Dimostrazione.* Per cominciare mostriamo che in  $X$  esiste uno schema di Cantor  $(U_s)_{s \in 2^{<\mathbb{N}}}$  tale che per ogni  $s \in 2^{<\mathbb{N}}$ :

- (1)  $U_s$  è aperto non vuoto.
- (2)  $diam(U_s) \leq 2^{-l(s)}$ .
- (3)  $\overline{U_{s \star i}} \subseteq U_s$ .

Procediamo ricorsivamente. Per  $s = \emptyset$  poniamo  $U_s$  un insieme qualsiasi che soddisfi (1) e (2), in particolare è lecito prendere una bolla di raggio  $1/2$ . Supponiamo ora di aver definito  $U_s$ . Essendo  $X$  perfetto esistono due punti distinti  $x_0, x_1 \in U_s$ . Siano  $U_{s \star 0}$  e  $U_{s \star 1}$  due interni aperti disgiunti di (rispettivamente)  $x_0$  e  $x_1$  di diametro al più  $2^{-l(s)-1}$  e con chiusura contenuta in  $U_s$ . Tali interni esistono: ogni spazio metrico è  $T_4$ , quindi  $x_0$  e  $x_1$  possono essere separati con interni disgiunti  $V_0$  e  $V_1$  contenuti in  $U_s$ . Questi possono essere raffinati ad interni  $W_0$  e  $W_1$  con chiusura contenuta in  $U_s$ . Per trovare gli interni desiderati basta allora intersecare  $W_0$  e  $W_1$  con  $B_{2^{-l(s)-2}}(x_i)$ .

$U_{s \star 0}$  e  $U_{s \star 1}$  soddisfano tutte le proprietà richieste: sono sottoinsiemi disgiunti di  $U_s$ , sono aperti non vuoti, hanno diametro sufficientemente piccolo e la loro chiusura è contenuta in  $U_s$ .

Costruiamo una mappa da  $\mathcal{C}$  in  $X$  nel modo seguente. Fissato  $x = (x_0, \dots) \in \mathcal{C}$  si consideri l'insieme  $S_x = \bigcap_{n \in \mathbb{N}} U_{x|n}$ , dove  $x|n = (x_0, \dots, x_n)$ . Per via della proprietà (3),  $S_x = \bigcap_{n \in \mathbb{N}} \overline{U_{x|n}}$ . Si consideri la successione  $(y_n)_{n \in \mathbb{N}}$ , con  $y_n \in U_{x|n}$  per ogni  $n$ . Grazie a (2) questa è una successione di Cauchy, per cui converge ad un punto  $y^*$ . Fissato  $k \in \mathbb{N}$ ,  $(y_n)_{n \in \mathbb{N}}$  è definitivamente a valori nel chiuso  $\overline{U_{x|k}}$ , dunque  $y^*$  è un elemento di  $\overline{U_{x|k}}$ . Per l'arbitrarietà di  $k$ ,  $y^* \in \bigcap_{n \in \mathbb{N}} \overline{U_{x|n}} = S_x$ . Concludiamo che  $S_x \neq \emptyset$ . Inoltre  $\text{diam}(S_x) = 0$ , quindi  $S_x$  è un singoletto. Chiamiamo  $f(x)$  l'unico elemento di  $S_x$ . La funzione  $f$  è iniettiva, poiché se  $x \neq y$  le due successioni  $\{U_{x|n}\}$  e  $\{U_{y|n}\}$  sono definitivamente disgiunte. Inoltre  $f$  è continua. Infatti se  $V$  è un qualunque intorno di  $f(x)$ , per questioni di diametro dovrà esistere  $N$  tale che  $U_{x|N} \subseteq V$ . Allora l'immagine dell'aperto  $\prod_{n \in \mathbb{N}} U_n$  con  $U_n = \{x_n\}$  per  $n \leq N$  e  $U_n = 2$  per  $n > N$  è contenuta in  $V$ . Essendo  $\mathcal{C}$  compatto e  $X$  di Hausdorff  $f$  è chiusa, quindi un omeomorfismo sull'immagine.  $\square$

**TEOREMA 21.** (Cantor-Bendixson) *Sia  $X$  uno spazio polacco, allora  $X$  si decompone in modo unico come  $X = P \cup N$ , dove  $P$  è un polacco perfetto e  $N$  è un aperto numerabile.*

*Dimostrazione.* Sia  $\mathcal{B}$  una base numerabile di  $X$ . Poniamo  $\mathcal{N} = \{B \in \mathcal{B} : |B| \leq \aleph_0\}$ . Sia  $N = \bigcup_{B \in \mathcal{N}} B$ .  $N$  è aperto poiché unione di aperti ed è numerabile perché unione numerabile di numerabili. Sia ora  $P = X \setminus N$ .  $P$  è chiuso in un polacco, quindi  $P$  è a sua volta polacco. Iniziamo mostrando che  $P$  è esattamente l'insieme dei punti di condensazione di  $X$ . Infatti un punto di condensazione non ha intorni numerabili, quindi non può appartenere a  $N$ , viceversa se un punto non appartiene a  $N$  allora non ha intorni della base numerabili e quindi non ha intorni numerabili in generale. Sia ora  $x \in P$  e sia  $U$  un suo intorno. Per definizione  $U$  è non numerabile, dunque anche  $U \cap P = U \setminus N$  è non numerabile, in particolare ha almeno due punti. Quindi  $P$  è perfetto.

Per mostrare l'unicità della decomposizione osserviamo anzitutto che dato uno spazio polacco perfetto  $Y$ , ogni suo punto è di condensazione. Infatti preso  $y \in Y$  e  $U$  intorno di  $y$ , esiste  $V$  intorno di  $y$  tale che  $\overline{V} \subseteq U$  (qui stiamo usando il fatto che  $X$  è T4).  $\overline{V}$  è chiuso in un polacco, dunque è esso stesso polacco. Inoltre  $\overline{V}$  è perfetto: preso un suo punto  $v$  e un suo intorno  $W$  ci sono due possibilità. O  $v \in V$  e quindi  $V \cap W$  è un intorno di  $v$ , per perfezione di  $Y$   $V \cap W$  deve contenere un altro punto oltre a  $v$ . Oppure  $v$  è di accumulazione per  $V$  e quindi  $W$  interseca  $V$  in un punto diverso da  $v$ . Dunque  $\overline{V}$  è un polacco perfetto, quindi  $|U| \geq |V| \geq 2^{\aleph_0}$ , di conseguenza  $y$  è di condensazione. Sia  $X = P' \cup N'$  una decomposizione con le proprietà dell'enunciato. Essendo  $N'$  aperto numerabile,  $N' \subseteq N$ . Invece  $P'$  è chiuso, dunque polacco. Ma  $P'$  è perfetto, dunque ogni suo punto è di condensazione, da cui  $P' \subseteq P$ . Si conclude che  $P = P'$  e  $N = N'$ .  $\square$

I risultati finora dimostrati ci consentono di dimostrare agevolmente l'ipotesi del continuo nel caso di spazi polacchi.

*Dimostrazione.* (Del teorema 16) Sia  $X$  uno spazio polacco con  $|X| > \aleph_0$ . Da 17 si ha che  $|X| \leq 2^{\aleph_0}$ . Decomponiamo grazie a 21:  $X = P \cup N$ . Essendo  $X$  non numerabile  $P \neq \emptyset$ , per cui da 20 si ha che  $|X| \geq 2^{\aleph_0}$ . Concludiamo che  $|X| = 2^{\aleph_0}$ .  $\square$

### Bibliografia

- [K] KENNETH KUNEN (2006), *Set theory: An introduction to independence proofs*. Elsevier
- [MF] MARTIN HILS E LOESER FRANÇOIS (2019), *A first journey through logic*. American mathematical society
- [M] MARCO MANETTI (2014), *Topologia*. Springer
- [Ke] ALEXANDER S. KECHRIS (1995), *Classical descriptive set theory*. Springer



# LA LEGGE DI BENFORD – OVVERO: COME HO IMPARATO AD EVITARE FRODI CON LA MATEMATICA

FILIPPO BERETTA

## 1. Introduzione

La teoria della Legge di Benford nasce dall'osservazione pratica di svariate successioni numeriche, provenienti da ambiti molto diversi tra loro. Dalla sua scoperta tramite lo studio delle tavole dei logaritmi, l'argomento ha conosciuto una crescita enorme. Ogni anno centinaia di articoli specifici vengono scritti da matematici, fisici, ingegneri, contabili e molti altri. Per procedere, si fissi per prima cosa la notazione. Ogni numero reale positivo maggiore di 1  $x$  può essere scritto come  $x = S_{10}(x) \cdot 10^k$ , con  $k$  intero, e  $S_{10}(x) \in [1; 10)$ . La Legge di Benford si concentra sulla distribuzione della prima cifra a partire da sinistra di un numero. Si noti che ovviamente tutta l'informazione su tale cifra è contenuta in  $S(x)$ . Tale approccio è utile per confrontare dataset con scale di grandezza molto diverse tra loro, concentrandosi su delle caratteristiche comuni.

L'intuizione potrebbe portare a supporre che ci sia circa l'11% di probabilità di osservare come cifra iniziale una quale in  $\{1, \dots, 9\}$ . In molti esempi di dataset ciò tuttavia non accade, ed anzi la probabilità di osservare una cifra  $d$  è approssimativamente  $\log_{10}(\frac{d+1}{d})$ . Vi è quindi circa il 30% di probabilità di osservare la cifra "1" e solo il 4.6% della cifra "9".

## 2. La storia

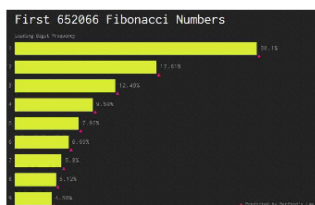
Anche se la legge prende il nome di Benford, il primo a segnalarne la presenza fu l'astronomo S. Newcomb, attorno al 1881. Studiando le tavole dei logaritmi in base 10, fu il primo ad evidenziare come le prime pagine di queste ultime si usurassero molto più delle ultime, segno di una distribuzione non uniforme della prima cifra dei numeri in esame. Newcomb fece anche una seconda osservazione cruciale, indicando espressamente la distribuzione probabilistica discreta della prima cifra: in particolare essa è "tale che le mantisse dei relativi logaritmi sono equiprobabili". Si ricorda che per mantissa si intende la parte fratta del logaritmo (in base 10) di un numero. Vedremo a breve come questa richiesta sia equivalente alla distribuzione di Benford.

Il nome di Frank Benford è invece legato alla notorietà che raggiunse il suo articolo "The Law of Anomalous Numbers" [B] del 1938, in cui vennero considerati 20 diversi dataset,

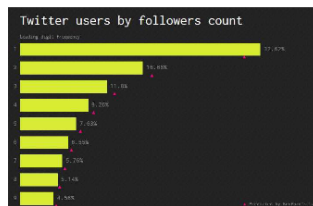
legati a fiumi, popolazioni, sequenze matematiche (come  $\sqrt{n}$ ,  $n!$ ,  $n^2$ , ...), sport eccetera, tutti approssimativamente caratterizzati dalla distribuzione che da lui avrebbe preso il nome. L'articolo presenta numerose osservazioni notevoli, per quanto solamente empiriche, ad esempio che un singolo dataset potrebbe non soddisfare la legge di Benford, ma l'unione di vari dataset diversi produce qualcosa di più vicino ad un "comportamento Benford".

Concludiamo questa sezione segnalando alcuni degli svariati esempi di dataset in cui emerge naturalmente la legge di Benford. Tra i più suggestivi vi è di certo l'uso che ne viene fatto in ambito di frodi bancarie. L'IRS (agenzia governativa americana deputata alla riscossione delle tasse) non fa mistero di usare questo strumento, e sono stati diversi gli scandali relativi a frodi scoperte tramite la Legge di Benford. Ciò accade ad esempio perché falsificando dei dati si tende a ricercare una equi-distribuzione troppo "perfetta", oppure a favorire alcune cifre specifiche, producendo così comportamenti ben lontani da quelli predetti dalla legge. Non mancano poi esempi di elezioni politiche, i cui risultati sono stati contestati facendo appello a delle distribuzioni non conformi alla legge di Benford. Tra gli esempi più famosi vi sono di certo quello delle elezioni in Iran del 2009 ed in America nel 2020.

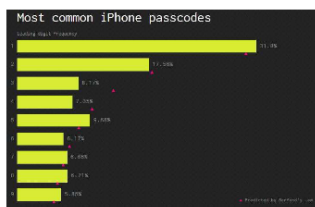
### Fibonacci numbers



### Twitter users by # followers



### Most common iPhone passcodes



### Distance of stars from Earth

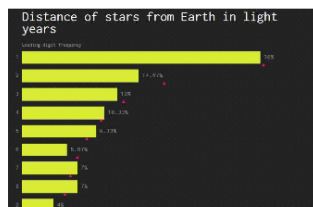


FIGURA 1. Esempi di Bias nella prima cifra

## 3. Formalizzazione della legge di Benford

Per cominciare, si definisce la distribuzione di probabilità discreta che prende il nome di Legge di Benford.

DEFINIZIONE 1. Un set di numeri soddisfa la Legge di Benford se la probabilità di osservare  $d \in \{1, \dots, 9\}$  come prima cifra è  $\log_{10}(\frac{d+1}{d})$ .

In generale questa definizione è difficile da verificare nella pratica, e ci concentreremo spesso su dataset che convergono verso una distribuzione Benford, aumentando il numero di dati in esame (si pensi ad esempio ad una successione numerica come  $\sqrt{n}$ ).

La prima osservazione, cruciale per quanto ovvia, è la seguente

OSSERVAZIONE 2. Dati due generici numeri positivi,  $\log_{10}(x) = \log_{10}(y) \bmod 1$  se e solo se  $S_{10}(y) = S_{10}(x)$ . Infatti  $\log_{10} x = \log_{10} S_{10}(x) + k$ . Per questo spesso si studia  $z = \log_{10}(x) \bmod 1$  per discutere la presenza della Legge di Benford.

Ispirati dalle discussioni di Newcomb, produciamo una seconda definizione. Vedremo in seguito l'equivalenza delle due.

DEFINIZIONE 3. Una successione  $\{y_n\}$  di numeri reali maggiori di 1 è equidistribuita mod 1 se, per ogni intervallo  $[a; b]$  contenuto in  $[0; 1]$ , la probabilità che  $y_n \bmod 1$  appartenga ad  $[a; b]$  tende a  $b - a$ :

$$\frac{\{n \leq N : y_n \bmod 1 \in [a; b]\}}{N} \rightarrow b - a$$

Si noti che questo equivale a richiedere la convergenza verso una distribuzione uniforme  $\mathcal{U}([0; 1])$ .

Riportiamo di seguito un teorema che, per quanto semplice, ci permetterà di fare alcune osservazioni interessanti nel seguito. La sua dimostrazione viene lasciata all'esercizio del lettore interessato.

TEOREMA 4. Per  $\beta \notin \mathbb{Q}$ , la successione numerica  $n\beta$  è equidistribuita mod 1.

La seguente osservazione permette di collegare i due concetti appena introdotti.

TEOREMA 5. Un set di dati reali positivi maggiori di 1  $\{x_i\}$  soddisfa la Legge di Benford in base 10 se e solo se  $\{y_i\}$  sono equidistribuiti mod 1, dove  $y_i = \log_{10} x_i$ .

*Dimostrazione.* Si vedrà solo una delle due implicazioni. Quella opposta è del tutto analoga, ripercorrendo le uguaglianze al contrario. Si osservi per prima cosa che l'insieme dei numeri con prima cifra  $d$ ,  $\{x_i : S(x_i) \in [d; d + 1)\}$ , coincide con  $\{x_i : \log_{10} S(x_i) \in [\log_{10} d; \log_{10}(d + 1))\}$ . Si ricordi inoltre che, come già osservato  $\log_{10} S(x_i) \equiv_1 \log_{10} x_i$ . Assumendo che la legge di Benford sussista per il set di dati  $\{x_i\}$ , la probabilità di questo evento è  $\log_{10}(\frac{d+1}{d}) = \log_{10}(d + 1) - \log_{10} d$ . Ma questa è anche la lunghezza dell'intervallo, una volta passati ai  $\log \bmod 1$ . Dunque gli  $y_i$  sono equidistribuiti.  $\square$

OSSERVAZIONE 6. Questa osservazione, molto semplice, ci permette discutere alcuni set di numeri che soddisfano certamente la legge di Benford. Tra questi,  $2^n$  è Benford in base 10, poichè  $\log_{10} 2 \notin \mathbb{Q}$ . Allo stesso modo, molte relazioni ricorsive forniscono dei dataset Benford, perché la soluzione della relativa equazione alle differenze è formata da potenze irrazionali. Ne è un esempio la successione dei numeri di Fibonacci,

descritti tramite la relazione  $a_{n+1} = a_n + a_{n-1}$ . La soluzione generale ha infatti forma:

$$a_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^n$$

In conclusione, è bene mettere in luce che la discussione da noi effettuata è stata limitata al caso di base 10, ma si può generalizzare in modo ovvio ad una generica base intera. In questo caso, la relazione fondamentale assume la seguente forma.

**TEOREMA 7.** *Un set di dati reali positivi maggiori di 1  $\{x_i\}$  soddisfa la Legge di Benford in base  $B$  se e solo se  $\{y_i\}$  sono equidistribuiti mod 1, dove  $y_i = \log_B x_i$ .*

#### 4. Insorgenza della legge di Benford in natura

Nelle precedenti sezioni è stata enunciata la legge di Benford, si è fornito un criterio pratico per caratterizzarla e se ne sono discussi alcuni esempi. Vorremmo ora mettere in luce il motivo dell'insorgenza in natura di questa distribuzione, apparentemente del tutto controintuitiva. Si vedranno prima alcune osservazioni euristiche, relative a dei casi specifici, e verrà in seguito enunciato un teorema limite, sulla falsariga del TLC.

**OSSERVAZIONE 8.** E' già stato osservato che le relazioni ricorsive, o più in generale le equazioni alle differenze, forniscono un pool di esempi di successioni che soddisfano la Legge di Benford. Dal momento che esse corrispondono ad una discretizzazione di equazioni differenziali ordinarie, che modellano molti processi fisici naturali, è naturale immaginare che la distribuzione Benford compaia in molti dataset presi dall'esperienza concreta.

**OSSERVAZIONE 9.** Tra gli esempi sopra menzionati, le progressioni geometriche godono di un ruolo di particolare interesse. In questo caso, è possibile fornire una dimostrazione diretta dell'insorgenza della legge di Benford. In particolare, per passare dal numero  $d$  a  $d + 1$ , si dovrà avere che  $d + 1 = k^{n_d} d$ , per cui

$$n_d = \frac{\log \frac{d+1}{d}}{\log k}$$

Allo stesso modo, per passare 1 a 10, si avrà  $n = \frac{\log 10}{\log k}$ . Semplificando, si ottiene  $\frac{n_d}{n} = \log_{10} \left( \frac{d+1}{d} \right)$

In ultimo, è opportuno tornare all'osservazione di Benford sul fatto che unioni di dataset siano più vicini ad una distribuzione Benford rispetto ai singoli. Prese allora delle variabili aleatorie a valori in  $[1; +\infty)$   $X_1, \dots, X_N$ , se ne può considerare il prodotto  $W_N := X_1 \cdot \dots \cdot X_N$ . Definite allora  $Y_i := \log_{10} X_i$ , è chiaro che  $V_N := \log_{10} W_N = Y_1 + \dots + Y_N$ . A quest'ultima somma di variabili aleatorie è possibile applicare il TLC per garantire la convergenza ad una v.a. Normale. La convergenza di  $W_N$  ad una distribuzione che sia Benford viene garantita dalla combinazione del TLC e del seguente:



TEOREMA 10. (Teorema Limite Centrale) Siano  $\{X_n\}$  variabili aleatorie i.i.d. ed in  $\mathcal{L}^2$ . Allora

$$S_N = \frac{X_1 + \dots + X_N - N\mu}{\sigma\sqrt{N}} \rightarrow N(0; 1) \text{ in legge}$$

Quindi approssimativamente, per  $N$  grande,

$$S_N \sim N(N\mu, N\sigma^2)$$

TEOREMA 11. Siano  $\{Y_n\}$  variabili aleatorie con le  $N(0, \sigma^2)$ . Allora per  $\sigma \rightarrow \infty$ ,

$$Y_n \text{ mod } 1 \rightarrow \mathcal{U}(0, 1) \text{ in legge}$$

Combinando i due teoremi si ottiene quindi la convergenza desiderata delle variabili  $W_N$  di partenza ad una distribuzione Benford. La dimostrazione del secondo risultato non viene riportata, essendo abbastanza tecnica, ma può essere trovata in [M2].

Concludiamo la trattazione riportando un teorema più forte del precedente, sempre tratto da [M2], che fornisce una condizione necessaria e sufficiente per la convergenza mod1 ad una uniforme. Per prima cosa, data una funzione  $f$  integrabile e supportata su  $[0; 1]$ , si consideri la sua trasformata di Fourier

$$\widehat{f}(n) := \int_0^1 e^{-2\pi i x n} f(x) dx$$

Vale allora il seguente risultato.

TEOREMA 12. (Nigrini, Miller, 2007) Siano  $\{Y_m\}$  variabili aleatorie indipendenti, supportate su  $[0; 1]$ , non necessariamente identicamente distribuite, con densità  $\{g_m\}$ . Una condizione necessaria e sufficiente perchè la somma  $Y_1 + \dots + Y_M \text{ mod } 1$  converga ad una  $\mathcal{U}(0; 1)$  in  $L^1([0, 1])$  è che per ogni  $n \neq 0$ ,  $\lim_{M \rightarrow \infty} \widehat{g}_1(n) \dots \widehat{g}_M(n) = 0$ . In particolare,  $W_N$  converge ad una distribuzione Benford base  $B$ , se  $Y_i = \log_B X_i$ .

### Bibliografia

- [B] F. BENFORD. *The Law of Anomalous Numbers*. Proceedings of the American Philosophical Society, 78: 551-572.
- [R] R. A. RAIMI. *The first digit problem*. Amer. Math. Monthly 83(7):521-538.
- [H] T. HILL. *The first-digit phenomenon*, American Scientist, 86:358-363.
- [M1] S. J. MILLER (2015) *Theory and Applications of Benford's Law*, Princeton University Press, Princeton,
- [M2] S. J. MILLER, M. NIGRINI *The Modulo 1 Central Limit Theorem and Benford's Law for Products*, International Journal of Algebra, 2(3): 119-130



# DA DIDONE A GROMOV: IL PROBLEMA ISOPERIMETRICO E LA SUA SOLUZIONE

BEATRICE OSTORERO VINCI

## Un problema dall'antichità

La prima descrizione in letteratura del problema isoperimetrico risale all'Eneide di Virgilio: nel racconto della fondazione di Cartagine, il poeta scrive che alla principessa Didone venne chiesto se volesse cedere tutta la sua fortuna al re locale in cambio della quantità di terra che sarebbe riuscita a recintare con la pelle di un bue. La principessa accettò la sfida e, presa la pelle del bue, la tagliò a strisce molto sottili; con queste realizzò un nastro molto lungo e lo utilizzò per recintare una porzione di terra di forma circolare. In questo modo Didone riuscì effettivamente ad isolare nella pelle del bue la massima area possibile e lì nacque Cartagine.

Letteratura a parte, il problema ha effettivamente origini molto antiche; infatti, era già noto al tempo dei greci, che però non riuscirono a risolverlo in modo rigoroso nella sua formulazione generale.

## Generalizzazione del problema

Esplicitamente il problema ha, in dimensione 2, la seguente formulazione:

Determinare la figura piana con area massima tra quelle aventi un perimetro di lunghezza fissata.

È possibile formulare l'analogo problema anche fissando l'area e chiedendo di determinare la figura con perimetro minimo tra quelle con area fissata.

Tra le due formulazioni non cambia molto, se non l'interpretazione che ne viene data o il contesto in cui si utilizzano, e possono essere associate alla stessa disuguaglianza, detta *disuguaglianza isoperimetrica*:

$$L^2 \geq 4\pi A$$

Dove  $L$  è la lunghezza del perimetro e  $A$  indica l'area della figura.

La soluzione al problema in due dimensioni è, in ogni caso, quella figura che satura tale disuguaglianza, e nel concreto risulta essere un cerchio. Infatti, detto  $r$  il suo raggio, si ha  $L^2 = (2\pi r)^2 = 4\pi^2 r^2$  e  $A = \pi r^2$ .

A noi però non basta sapere quale sia la soluzione in questo caso particolare, ma vogliamo risolvere il quesito in una dimensione  $n$  generica.

Per fare questo, la prima cosa di cui dobbiamo preoccuparci e di capire cosa sia effettivamente un perimetro in dimensione maggiore di 2.

Consideriamo quindi  $K \subset \mathbb{R}^n$  aperto, limitato, convesso e tale che l'origine degli assi vi appartenga, e utilizziamo questo aperto per definire la seguente norma, pesata rispetto alla direzione:

$$\|v\|_* := \sup\{x \cdot v : x \in K\}, \quad v \in \mathbb{S}^{n-1}$$

Dove con  $\mathbb{S}^{n-1}$  stiamo indicando la sfera di  $\mathbb{R}^n$ .

Sia poi  $E \subset \mathbb{R}^n$  aperto, con bordo  $\partial E$  liscio o poliedrico e orientato da  $v_E$ .

DEFINIZIONE 1. Definiamo perimetro anisotropico di  $E$

$$P_K(E) := \int_{\partial E} \|v_E\|_* d\mathcal{H}^{n-1}(x)$$

Con  $\mathcal{H}^{n-1}$  misura di Hausdorff  $(n-1)$ -dimensionale.

Questa definizione può, a colpo d'occhio, sembrare distante da quella che intuitivamente daremmo di perimetro cercando di generalizzare quella relativa ad  $\mathbb{R}^2$ .

In effetti, per un insieme  $K$  generico, i.e. non dotato di particolari simmetrie,  $P_K(E)$  non ha le stesse proprietà del perimetro euclideo: in particolare, non risulta invariante per trasformazioni legate a matrici ortogonali  $O(n)$  (roto-omotetie), o speciali ortogonali  $SO(n)$  (rotazioni).

Mantiene però l'invarianza per traslazioni e, se si considera  $K = B$  bolla unitaria  $n$ -dimensionale, la definizione coincide con quella di perimetro euclideo; infatti ritroviamo:

$$\|v_E\|_* = 1 \text{ da cui } P_B(E) = \int_{\partial E} d\mathcal{H}^{n-1}(x)$$

Invece di risolvere il problema per un insieme  $K$  generico, prendiamo in esame quest'ultimo caso e, come già fatto per  $n = 2$ , scriviamo la disuguaglianza associata:

$$P_B(E) \geq n|B|^{\frac{1}{n}}|E|^{\frac{n-1}{n}}. \quad (1)$$

A livello di notazione, segnaliamo che  $|X|$  indica la misura  $n$ -dimensionale di  $X \subset \mathbb{R}^n$ .

Si vuole quindi dimostrare la disuguaglianza (1) e capire per quali insiemi risulta effettivamente saturata.

### Alcuni strumenti utili

Allo scopo di arrivare all'obiettivo che ci siamo posti, introduciamo alcuni elementi teorici che verranno utilizzati in seguito.

### Teoria della misura

Il primo strumento che introduciamo consiste in qualche elemento di teoria della misura.

Innanzitutto ricordiamo che dato  $(X, d)$  uno spazio metrico, possiamo indicare con  $\mathcal{B}(X)$  la  $\sigma$ -algebra di Borel su  $X$ , e che data una misura di Borel  $\mu: \mathcal{B}(X) \rightarrow [0, +\infty]$ , diciamo che questa è:

- finita se  $\mu(X) < +\infty$ ;
- di probabilità se è finita e  $\mu(X) = 1$ .

Inoltre, risulta particolarmente rilevante la seguente definizione:

**DEFINIZIONE 2 (Misura immagine).** Siano  $X, Y$  spazi metrici,  $\mu$  una misura di probabilità su  $X$ , e sia  $T: X \rightarrow Y$  misurabile.

Si definisce misura immagine di  $\mu$  tramite la mappa  $T$  la misura  $\nu := T_{\#}\mu$ , tale che  $\forall E \in \mathcal{B}(Y) \nu(E) = \mu(T^{-1}(E))$ .

Una proprietà della misura immagine che sarà utile ai fini della dimostrazione è il *cambio di variabile*.

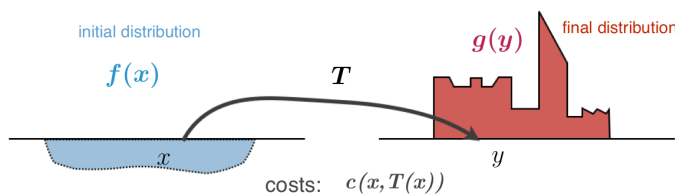
**PROPOSIZIONE (Cambio di variabile).** Dati  $X, Y$  spazi metrici,  $T: X \rightarrow Y$  una mappa misurabile e  $f \in L^1(Y, T_{\#}\mu)$ , vale

$$\int_Y f(y) d(T_{\#}\mu)(y) = \int_X f(T(x)) d\mu(x).$$

### La teoria del trasporto ottimo

L'altro strumento fondamentale sarà la teoria del trasporto ottimo. Ma che cos'è?

Per introdurla intuitivamente, possiamo pensare di avere come obiettivo quello di trasferire una certa distribuzione di massa da un luogo ad un altro, minimizzando il costo del trasporto. Per esempio, possiamo pensare di voler portare una certa quantità di sabbia che si trova in una configurazione descritta dalla funzione  $f$  come un figura, in una nuova configurazione, descritta da una seconda funzione  $g$ .



La mappa  $T: X \rightarrow Y$  che realizza tale trasferimento prende il nome di *mappa di trasporto*, e, come idea, possiamo pensare alla funzione costo associata  $c(x, T(x))$  come a qualcosa di legato al lavoro che si compie per il trasporto della massa.

Il problema, a livello matematico, a cui tale teoria cerca di dare una risposta venne formulato per la prima volta da Monge nel 1781 ed è il seguente:

Date  $\mu, \nu$  misure di probabilità definite su  $X$  e su  $Y$  rispettivamente, e considerato un funzionale  $\mathbb{M}(T) := \int_X c(x, T(x)) d\mu(x)$ , determinare

$$\inf_{T: \nu = T\#\mu} \mathbb{M}(T)$$

e la mappa di trasporto  $T$  che lo realizza.

La funzione  $T$  che risolve il problema è detta *mappa di trasporto ottimale* e siamo interessati nello specifico a capire se e sotto quali ipotesi tale funzione esiste e di quali proprietà gode.

In particolare, per l'esistenza ci viene in aiuto Brenier.

Il setting in cui ci poniamo è  $\mathbb{R}^n$  e consideriamo  $X, Y$  due sottoinsiemi. Inoltre, citiamo a titolo di completezza (anche se per quel che interessa a noi del risultato non ha alcun ruolo), che la funzione costo utilizzata è  $c(x, y) = \frac{1}{2}|x - y|^2$ . Allora:

TEOREMA (Brenier). Siano  $X, Y \subseteq \mathbb{R}^n$ ,

$$\mu \in L^2(X) = \{\mu \text{ misura di probabilità su } X : \int_X |x|^2 d\mu(x) < \infty\},$$

$$\nu \in L^2(Y) = \{\nu \text{ misura di probabilità su } Y : \int_Y |y|^2 d\nu(y) < \infty\}.$$

Supponiamo  $\mu(\{x\}) = 0 \forall x \in X$ . Allora:

- (1) Esiste un unico piano di trasporto ottimale  $\pi$  tale che  $\pi = (Id \times T)\#\mu$ , dove  $T := \nabla\varphi$  è l'unico gradiente di una funzione convessa definito  $\mu$ -quasi ovunque tale che  $T\#\mu = \nu$ .
- (2)  $T$  è l'unica soluzione del problema di Monge con costo quadratico.

Quindi, quantomeno in queste circostanze, sappiamo che il problema di Monge ammette soluzione.

Prima di guardare alle proprietà della mappa  $T$ , cerchiamo di capire cosa c'entra tutto questo con il nostro problema iniziale.

Supponiamo di considerare  $X = E$  e  $Y = B$  e poniamo poi  $f(x) := \frac{\chi_E(x)}{|E|}$  e  $g(y) := \frac{\chi_B(y)}{|B|}$ . A partire da queste, possiamo definire una misura su  $E$  e una su  $B$ , rispettivamente come  $|F|_E = \int_F f(x) dx, \forall F \subset E$  in  $\mathcal{B}(E)$ , e  $|C|_B = \int_C g(x) dx, \forall C \subset B$  in  $\mathcal{B}(B)$ .

Si osservi che proprio per come sono definite, queste risultano essere misure di probabilità sui rispettivi spazi; inoltre, dal fatto che  $E$  e  $B$  sono limitati, segue anche che  $f \in L^2(E)$ ,  $g \in L^2(B)$ . Conseguentemente, si possono sostituire le misure  $\mu$  e  $\nu$  che compaiono nell'enunciato del teorema di Brenier con  $f$  e  $g$  rispettivamente.

Quello che possiamo concludere, proprio applicando il teorema, è che esiste una mappa  $T: E \rightarrow B$  che "trasporta"  $f$  in  $g$ .

Adesso che in qualche modo abbiamo ricondotto tutto quello che è stato introdotto sulla teoria del trasporto ottimo al problema isoperimetrico, notiamo che  $T$  presenta tre proprietà interessanti.

PROPOSIZIONE (Proprietà 1).  $T(E) \subseteq B$ , e quindi  $\|T\| \leq 1$  su  $\partial E$ .

*Dimostrazione.* Cominciamo a provare che l'immagine di  $E$  tramite  $T$  è interamente contenuta in  $B$ . Supponiamo per assurdo che ci sia  $F \subseteq E$  di misura non nulla, tale

che  $T(F) \cap B = \emptyset$ . Allora

$$0 = |T(F)|_B = \int_{T(F)} dg(y) = \int_F df(x) = \frac{|F|}{|E|}$$

Ma questo è assurdo per il fatto che  $F$  non ha misura nulla.

Quindi, se prendiamo un punto  $x \in E$ ,  $T(x) \in B$  e di conseguenza  $\|T(x)\| < 1$ . Ma allora su  $\partial E$ , essendo  $T$  continua, deve essere tale che  $\|T\| \leq 1$ , il che prova anche la seconda parte della tesi.  $\square$

PROPOSIZIONE (Proprietà 2).  $|\det(JT_x)| = \frac{|B|}{|E|} \quad \forall x \in E$ .

Si osservi che abbiamo indicato con  $JT_x$  la matrice Jacobiana di  $T$  valutata in  $x$ .

*Dimostrazione.* Sia  $F \subset E$ , allora:

$$\frac{|F|}{|E|} = \int_F \frac{\chi_B(T(x))}{|B|} \det(JT_x) dx \tag{2}$$

$$= \int_{T(F)} \frac{\chi_B(y)}{|B|} dy \tag{3}$$

$$= \int_{T(F)} \frac{1}{|B|} dy \tag{4}$$

$$= \frac{|T(F)|}{|B|} \tag{5}$$

Dove:

- Nel passaggio da (1) a (2) è stato usato l'usuale cambio di variabili  $y = T(x)$ ;
- Nel passare da (2) a (3), abbiamo sfruttato il fatto che per la proprietà precedente  $T(F) \subset B$ .

Da  $\frac{|F|}{|E|} = \frac{|T(F)|}{|B|}$  segue che  $\frac{|B|}{|E|} = \frac{|T(F)|}{|F|}$ ,  $\forall F \in \mathcal{B}(\mathbb{R}^n)$ . Quindi in particolare, per  $F = B_r(x)$ ,  $x \in E$ ,  $r$  opportuno, si ha:

$$\lim_{r \rightarrow 0} \frac{|T(F)|}{|F|} = |\det(JT_x)|.$$

$\square$

Infine, l'ultima proprietà interessante della mappa  $T$  coinvolge la sua divergenza:

PROPOSIZIONE (Proprietà 3).  $Div(T) \geq n(\det(JT))^{1/n}$ .

*Dimostrazione.* Per dimostrare questa proprietà, si ricordi che dal Teorema di Brenier  $T = \nabla\varphi$ , con  $\varphi$  una funzione convessa.

Consideriamo quindi:

$$Div(T) = \sum_{i=1}^n \frac{\partial T_i}{\partial x_i} = \Delta\varphi = Tr(H\varphi) = \sum_{i=1}^n \lambda_i,$$

Dove  $H\varphi$  è la matrice Hessiana di  $\varphi$  e  $\lambda_i$  sono i suoi autovalori.

Per la disuguaglianza aritmetico-geometrica si ha:

$$\operatorname{Div}(T) = n \left( \frac{1}{n} \sum_{i=1}^n \lambda_i \right) \geq n \left( \prod_{i=1}^n \lambda_i \right)^{\frac{1}{n}} = n(\det(JT))^{\frac{1}{n}}.$$

Si noti che, essendo  $\varphi$  convessa,  $H\varphi$  è semidefinita positiva e di conseguenza  $\lambda_i \geq 0 \forall i$ ; pertanto la scrittura  $(\prod_{i=1}^n \lambda_i)^{\frac{1}{n}}$  ha senso.  $\square$

### Risoluzione del problema

Abbiamo quindi ora a disposizione tutti gli strumenti necessari per dimostrare la disuguaglianza 1, che ricordiamo essere:

$$P_B(E) \geq n|B|^{\frac{1}{n}}|E|^{\frac{n-1}{n}}.$$

*Dimostrazione.* Per definizione  $P(E) = \int_{\partial E} 1 d\mathcal{H}^{n-1}(x)$ , e per la Proprietà 1  $\|T\| \leq 1$  su  $\partial E$ . Quindi  $P(E) \geq \int_{\partial E} \|T\| d\mathcal{H}^{n-1}(x)$ .

Per Cauchy-Schwartz  $\|T\| \geq T \cdot \nu_E$ , e applicando il Teorema della Divergenza, si ottiene

$$P(E) \geq \int_{\partial E} T \cdot \nu_E d\mathcal{H}^{n-1}(x) = \int_E \operatorname{Div}(T) d\mathcal{H}^n(x).$$

A questo punto, dalla Proprietà 3 si ha  $\int_E \operatorname{Div}(T) d\mathcal{H}^n(x) \geq \int_E n(\det(JT))^{\frac{1}{n}} d\mathcal{H}^n(x)$ , ed essendo  $\det(JT) = \frac{|B|}{|E|}$ , ricaviamo:

$$P(E) \geq n \int_E \left( \frac{|B|}{|E|} \right)^{\frac{1}{n}} d\mathcal{H}^n(x) = n|B|^{\frac{1}{n}}|E|^{\frac{n-1}{n}}.$$

$\square$

### Saturazione della disuguaglianza

Per ottenere delle informazioni sulla figura che risolve il problema e sulla sua misura, vogliamo che la disuguaglianza  $P(E) \geq n|B|^{\frac{1}{n}}|E|^{\frac{n-1}{n}}$  sia saturata. Pertanto, tutte le disuguaglianze scritte nella dimostrazione precedente devono essere saturate, e questo forza  $T$  a soddisfare condizioni ben precise:

- (1)  $\int_{\partial E} 1 d\mathcal{H}^{n-1}(x) \geq \int_{\partial E} |T| \rightsquigarrow |T| = 1$  su  $\partial E$ ;
- (2)  $\int_{\partial E} |T| \geq \int_{\partial E} T \cdot \nu_E \rightsquigarrow T \parallel \nu_E$  e concordi;
- (3)  $\int_E \operatorname{Div}(T) \geq \int_E n(\det(JT))^{\frac{1}{n}}$

$$\begin{aligned} \rightsquigarrow \frac{1}{n} \sum_{i=1}^n \lambda_i &= \left( \prod_{i=1}^n \lambda_i \right)^{\frac{1}{n}} \iff \lambda_i = c \forall i \\ &\iff JT = cI \text{ in ogni punto di } E \\ &\iff T(x) = c\underline{x} + T(\mathbf{0}) \end{aligned}$$

Con  $T(\mathbf{0})$  una costante opportuna che non è necessario determinare esplicitamente.



Segue che per  $y \in B$ , la mappa  $T^{-1}$  manda  $y \mapsto \frac{y-T(0)}{c} \in E$ , ossia agisce su  $B$  come una combinazione di una traslazione e una dilatazione.

Ma allora, essendo  $B$  una bolla  $n$ -dimensionale, l'insieme  $E = T^{-1}(B)$  che satura la disuguaglianza, è necessariamente una bolla  $n$ -dimensionale!

### Bibliografia

- [F] FIGALLI, A., MAGGI, F. & PRATELLI, A. (2010) *A mass transportation approach to quantitative isoperimetric inequalities*  
Invent. math. 182, 167–211. <https://doi.org/10.1007/s00222-010-0261-z>



# DIMOSTRAZIONI PROBABILISTICHE DI RISULTATI ANALITICI

GABRIELE CASSESE

È indubbio che alcuni risultati di analisi servano come fondamenta della teoria della probabilità e del calcolo stocastico. Tuttavia, un rapporto tra due così estese branche della matematica non può essere monodirezionale; giunge quindi il momento di chiedersi non cosa possa fare l'analisi per la probabilità, ma cosa possa fare la probabilità per l'analisi. Analizzeremo due diverse applicazioni all'analisi di tecniche probabilistiche:

- (1) All'analisi funzionale: buona parte di queste pagine, nonché la totalità del seminario, è stata dedicata allo studio della proprietà di Radon-Nykodim in spazi di Banach. Come vedremo, la teoria delle martingale discrete si rivela estremamente feconda in questo contesto
- (2) All'analisi "complessa": nell'ultima sezione, non discussa nel seminario, si accennano brevemente alcune interessanti applicazioni del calcolo stocastico all'analisi complessa elementare.

## 1. Radon-Nikodym: un breve ripasso

**DEFINIZIONE 1.** Sia  $(X, \Sigma)$  uno spazio misurabile e  $\mu, \nu$  due misure definite su tale spazio. Diciamo che  $\mu$  è assolutamente continua rispetto a  $\nu$  ( $\mu \ll \nu$ ) se, per ogni  $A \in \Sigma$ ,  $\nu(A) = 0$  implica  $\mu(A) = 0$ . Diciamo che  $\mu \perp \nu$  ( $\mu$  e  $\nu$  sono singolari) se esiste  $A \in \Sigma$  :  $\mu(A) = 0$ ;  $\nu(X \setminus A) = 0$ .

**TEOREMA 2** (Lebesgue-Radon-Nikodym). *Sia  $(X, \Sigma)$  uno spazio misurabile e siano  $\mu, \nu$  due misure  $\sigma$ -finite<sup>1</sup> su tale spazio. Esiste una funzione misurabile  $f \geq 0$  ed una misura  $\sigma \perp \nu$  tale che per ogni  $A \in \Sigma$  si ha*

$$\mu(A) = \int_A f d\nu + \sigma(A). \quad (1)$$

*Dimostrazione.* Supponiamo che le misure siano di probabilità<sup>2</sup>. Sia  $H$  lo spazio di Hilbert  $L^2(\mu + \nu)$  e sia  $T$  il funzionale  $T(f) := \int f d\mu$ . Essendo  $T$  continuo, esiste  $g \in H$

---

<sup>1</sup>È in realtà sufficiente richiedere che  $\mu + \nu$  sia localizzabile ma analizzare questa estensione ci porterebbe su una strada completamente diversa

<sup>2</sup>l'estensione al caso  $\sigma$ -finito è lasciata come esercizio

tale che

$$T(f) = \int f g d(\mu + \nu).$$

Una semplice manipolazione algebrica ci porta ad ottenere

$$\int f(1-g)d\mu = \int f g d\nu.$$

A questo punto è evidente che  $0 \leq g \leq 1$   $\mu + \nu$ -quasi ovunque. Sia  $A = \{g = 1\}$  e siano  $\sigma := \mu|_A$ ,  $\tilde{\mu} := \mu - \sigma$ . Evidentemente  $\sigma \perp \nu$ . Per costruzione,  $0 \leq g < 1$   $\tilde{\mu} + \nu$  quasi ovunque e quindi per il teorema di convergenza monotona se  $f \geq 0$

$$\int f d\tilde{\mu} = \int \sum f g^n (1-g) d\tilde{\mu} = \int \sum f g^n d\nu = \int f \frac{1}{1-g} d\nu.$$

Definita  $h := \frac{1}{1-g}$ , ne segue che per ogni  $f \in L^2(\mu + \nu)$ ,  $f \geq 0$  si ha

$$\int f d\mu = \int f d\tilde{\mu} + \int f d\sigma = \int f h d\nu + \int f d\sigma.$$

□

## 2. Misure a valori vettoriali e RNP

**DEFINIZIONE 3.** Sia  $V$  uno spazio vettoriale topologico e sia  $(X, \Sigma)$  uno spazio misurabili. Una funzione  $\mu : \Sigma \rightarrow V$  si dice misura a valori in  $V$  se data una successione di elementi di  $\Sigma$  disgiunti  $\mu(\bigsqcup A_n) = \sum_n \mu(A_n)$ , dove la somma va interpretata come convergente nella topologia di  $V$ . Come per le misure con segno, si definisce la variazione totale di  $\mu$ , denotata  $\|\mu\|$  mediante  $\|\mu\|(A) = \sup_{\mathcal{F} \in \mathcal{S}(A)} \sum_{E \in \mathcal{F}} \|\mu(E)\|$ , dove  $\mathcal{S}(A)$  è l'insieme di partizioni misurabili finite di  $A$ <sup>3</sup>.

**ESEMPIO 4.** Sia  $(X, \sigma) = ([0, 1], \mathcal{B})$ ,  $V = c_0$  e  $\mu(A) = \left( \int_A \sin(2\pi n x) dx \right)_n$  (il fatto che  $\mu(A)$  appartenga a  $c_0$  per ogni  $A$  è una conseguenza del lemma di Riemann-Lebesgue): è facile vedere che  $\mu$  è una misura a valori vettoriali.

**ESEMPIO 5.** Una misura a valori di proiettore, fondamentale nella formulazione più generale del teorema spettrale, è una misura a valori vettoriali se  $B(H)$  viene equipaggiato della strong operator topology.

È inoltre rilevante osservare che, contrariamente a quanto succede nel caso finito-dimensionale, le misure limitate non hanno necessariamente variazione limitata<sup>4</sup>:

**ESEMPIO 6.**  $(X, \sigma) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ ,  $\mu : \mathcal{P}(\mathbb{N}) \rightarrow \ell_\infty$ ;  $\mu(\{n\}) = \frac{1}{n} e_n$ , dove  $e_k := (\delta_k^n)_n$ . Seppur  $\|\mu(\{A\})\| \leq 1$ , è facile vedere che  $\|\mu\|(\mathbb{N}) = \sum \frac{1}{n} = \infty$ .

D'ora in poi supporremo che ogni misura a valori vettoriali abbia variazione finita. Necessitiamo di un'altra generalizzazione infinito-dimensionale di un concetto

<sup>3</sup>che  $\|\mu\|$  sia una misura non è ovvio ed è lasciato come spunto di riflessione al lettore interessato

<sup>4</sup>il problema fondamentale è la mancanza di compattezza della bolla unitaria

usualmente incontrato nel caso monodimensionale: l'integrale di funzioni a valori vettoriali.

DEFINIZIONE 7. Sia  $(X, \Sigma, \mu)$  uno spazio misurabile,  $V$  uno spazio di Banach e  $f : X \rightarrow V$  una funzione. Essa è detta

- (1) fortemente misurabile se esiste una successione di funzioni semplici misurabili  $s_n : X \rightarrow V$  tali che per ogni  $x$  si ha  $f(x) - s_n(x) \rightarrow 0$ .
- (2) misurabile se è misurabile come mappa tra spazi misurabili  $(X, \Sigma) \rightarrow (V, \mathcal{B}(V))$ .
- (3) debolmente misurabile se per ogni  $\varphi \in V'$  si ha che  $\varphi \circ f$  è misurabile

Se  $f$  è fortemente misurabile, diciamo che essa è integrabile secondo Bochner se  $\lim \int \|f - s_n\| d\mu = 0$ . In tal caso, definiamo  $\int f d\mu := \lim \int s_n$ .

OSSERVAZIONE 8. È evidente che quasi tutte le definizioni precedenti hanno senso anche in un generico spazio di Frechet: l'unica non ovvia da estendere è l'integrabilità secondo Bochner; diciamo che  $f$  è integrabile secondo Bochner se  $\lim \int d(f, s_n) = 0$  e in tal caso  $\int f := \lim \int s_n$ . È possibile costruire differenti nozioni di integrale che valgono in contesti più generali <sup>5</sup> ma anche lì è necessaria qualche ipotesi di completezza (generalmente, quasi-completezza) dello spazio, senza la quale non si può garantire l'esistenza dell'integrale

È un facile esercizio dimostrare

PROPOSIZIONE 9. Sia  $f : (X, \sigma, \mu) \rightarrow V$  fortemente misurabile.  $f$  è Bochner integrabile sse  $\int \|f(x)\| d\mu < \infty$ .

Possiamo dunque definire lo spazio  $L^p((X, \Sigma, \mu), V)$  come ovvio, equipaggiandolo della norma  $\|f\|_p = \left(\int_X \|f\|^p d\mu\right)^{\frac{1}{p}}$ . Come ovvio, è molto utile avere un criterio rapido per determinare la misurabilità forte di funzioni: tale richiesta è soddisfatta dal criterio di Pettis:

TEOREMA 10. Sia  $f : (X, \Sigma) \rightarrow V$ . Essa è fortemente misurabile sse è debolmente misurabile ed ha range separabile, i.e.  $f(X)$  è separabile.

*Dimostrazione.* Se  $f$  è fortemente misurabile è ovviamente debolmente misurabile; che abbia immagine separabile si dimostra osservando che  $\overline{\text{span}}(\cup s_n(X)) \supset f(X)$  e che il termine a sinistra della precedente inclusione è separabile. Per dimostrare l'altra direzione, supponiamo senza perdita di generalità che  $V$  sia separabile. Sia  $\{e_n\}$  il sottoinsieme separabile e sia  $\{\varphi_k\}$  la famiglia di funzioni da  $V$  in se stesso tali che per ogni  $k$

$$\varphi_k(x) \in \{e_1, \dots, e_k\}$$

$$\|x - \varphi_k(x)\| = \min_{1 \leq i \leq k} \|x - e_i\|$$

<sup>5</sup>e.g. l'integrale di Dunford-Pettis  $\int f d\mu$  è definito come qualunque vettore  $v$  tale per cui per ogni funzionale continuo  $\varphi$  si ha  $\varphi(v) = \int \varphi(f) d\mu$ .

Per costruzione,  $\varphi_n \circ f$  è una famiglia di funzioni semplici che converge puntualmente a  $f$ . Rimane da dimostrare che  $\varphi_n \circ f$  è misurabile. Per dimostrarlo osserviamo che

$$\varphi_n \circ f^{-1}(e_k) = \left\{ x : \|f(x) - e_k\| = \min_{1 \leq i \leq n} \|f(x) - e_i\| \right\}.$$

Per separabilità di  $V$  esiste  $\mathcal{F} \subset V'$  numerabile denso in  $(B_{V'}, \sigma(V', V))$ ; la funzione  $\|f(x) - e_k\|$  è quindi  $\|f(x) - e_k\| = \sup_{y \in \mathcal{F}} y(f(x) - e_k)$  ed è quindi misurabile. Allo stesso modo si dimostra che  $\min_{1 \leq i \leq n} \|f(x) - e_i\|$  è misurabile e il risultato segue.  $\square$

Visti queste definizioni, siamo ora in grado di formalizzare il teorema di Radon-Nykodim in un contesto più generale:

**DEFINIZIONE 11.** Sia  $V$  uno spazio di Banach. Esso ha la proprietà di Radon-Nikodym (RNP) se data una misura  $\mu$  a valori in  $V$  e  $\nu$  misura finita tali che  $\|\mu\| \ll \nu^6$  esiste  $f \in L^1(\nu)$  tale che per ogni  $A$  misurabile vale

$$\mu(A) = \int_A f d\nu$$

**PROPOSIZIONE 12.**  $c_0$  non ha la proprietà di Radon-Nykodim

*Dimostrazione.* Sia  $\mu$  la misura a valori in  $c_0$  costruita nell'esempio precedente. Questa è evidentemente a variazione limitata e  $\mu \ll \lambda$ , dove  $\lambda$  è la misura di Lebesgue su  $[0, 1]$ . Rimane da dimostrare che non esiste  $f$  a valori in  $c_0$  tale che per ogni  $A$

$$\mu(A) = \int_A f d\lambda$$

Supponiamo per assurdo che tale funzione esista. Evidentemente  $f(x)$  deve coincidere con  $h := (\sin(2\pi nx))_n$ ; è sufficiente dimostrare che  $h \notin L^1([0, 1], c_0)$  per ottenere l'assurdo. Per farlo, sia  $A_n := \{x \in [0, 1] : \sin^2(nx) \geq \frac{1}{2}\}$ . Un rapido calcolo dimostra  $\lambda(A_n) = \frac{1}{4}$  per ogni  $n$  e dunque  $\mu(\limsup A_n) \geq \limsup \mu(A_n) = \frac{1}{4}$ . Tuttavia, preso  $x \in \limsup A_n$  abbiamo  $h(x) \notin c_0$ , dando la contraddizione cercata.  $\square$

### 3. Caratterizzazione di RNP

È nel tentativo di determinare precisamente quali spazi abbiano la RNP che probabilità ci viene in aiuto. Ricordiamo che nel caso scalare, dal teorema di Radon-Nikodym si ottiene facilmente il teorema di convergenza di Doob: una martingala discreta uniformemente integrabile converge in norma  $L^1$  e quasi ovunque ad una variabile aleatoria. Viceversa, è possibile ottenere dal teorema di Doob il teorema di Radon-Nikodym; è possibile che si sia una relazione simile in uno spazio di Banach? Per capirlo dobbiamo innanzitutto definire il concetto di valore atteso condizionato per variabili aleatorie a valori in spazi di Banach; in quanto segue identifichiamo le funzioni semplici  $s = \sum 1_{A_n} x_n$  con gli elementi del prodotto tensoriale  $L^1(\Omega) \otimes X$ ,  $\tilde{s} = \sum 1_{A_n} \otimes x_n$ .

<sup>6</sup>per brevità d'ora in poi scriveremo direttamente  $\mu \ll \nu$

**PROPOSIZIONE 13.** *Dato uno spazio di probabilità  $(\Omega, \Sigma, \mathbb{P})$ , una  $\sigma$ -algebra  $\mathcal{F} \subset \Sigma$  ed uno spazio di Banach  $X$ , l'operatore  $\mathbb{E}^{\mathcal{F}} := \mathbb{E}^{\mathcal{F}} \otimes I$  definito su  $L^1(\Omega) \otimes X$  ha norma 1 ed ammette una ed una sola estensione a  $L^1(\Omega, X)$ <sup>7</sup>.*

*Dimostrazione.* Sia  $s \in L^1(\Omega) \otimes X, s = \sum 1_{A_n} \otimes x_n$ . Ovviamente  $\|\mathbb{E}^{\mathcal{F}} \otimes I\| \geq 1$ . D'altro canto,

$$\|\mathbb{E}^{\mathcal{F}} \otimes I(\sum 1_{A_n} x_n)\| \leq \sum \mathbb{E}^{\mathcal{F}}(1_{A_n}) \|x_n\| \leq \sum \mathbb{P}(A_n) \|x_n\| = \|s\|_{L^1(\Omega, X)}$$

Il risultato segue ora dalla densità delle funzioni semplici in  $L^1(\Omega, X)$ . □

Dalla definizione seguono due utili proprietà:  $\|\mathbb{E}(X|\mathcal{F})\| \leq \mathbb{E}(\|X\||\mathcal{F})$  e preso  $\varphi$  funzionale lineare continuo,  $\mathbb{E}(\varphi(X)|\mathcal{F}) = \varphi(\mathbb{E}(X|\mathcal{F}))$ . La definizione di martingala è ora identica a quella scalare.

**DEFINIZIONE 14.** Sia  $X$  uno spazio di Banach. Dato  $p \in [1, \infty]$ , si dice che  $X$  ha la proprietà di  $p$ -convergenza di martingale ( $p$ -MCP) se per ogni spazio di probabilità  $(\Omega, \Sigma, \mathbb{P})$  e una martingala  $(X_n, \mathcal{F}_n)$  a valori in  $X$

- (1)  $p = 1$ : l'uniforme integrabilità della martingala e la limitatezza in  $L^1$  implicano la convergenza in norma  $L^1$  e quasi certamente ad una variabile aleatoria  $\tilde{X}$
- (2)  $p \in (1, \infty)$ : la  $L^p$  limitatezza della martingala implica la convergenza quasi quasi certa ed in  $p$ -norma ad una variabile aleatoria  $\tilde{X} \in L^p$ .
- (3)  $p = \infty$ : la  $L^\infty$  limitatezza della martingala implica la convergenza quasi certa ad una variabile  $\tilde{X} \in L^\infty$ .

**OSSERVAZIONE 15.** La 1-MCP implica la  $p$ -MCP per  $p > 1$ : presa una martingala  $X_n$   $L^p$ -limitata questa è necessariamente uniformemente integrabile, quindi converge q.c. ad una  $\tilde{X} \in L^p$  grazie al lemma di Fatou.  $\|X_n - \tilde{X}\|$  è dunque una submartingala uniformemente limitata in  $L^p$ ; la disuguaglianza di Doob implica

$$\left\| \sup_n \|X_n - \tilde{X}\| \right\|_p < \infty.$$

Il DCT implica ora che  $\mathbb{E}(\|X_n - \tilde{X}\|^p) \rightarrow 0$ , dando la convergenza richiesta. Il caso  $p = \infty$  è banale.

**TEOREMA 16.** *Sia  $X$  uno spazio di Banach. Sono equivalenti le proprietà:*

- (1)  $X$  ha la RNP
- (2)  $X$  ha la 1-MCP
- (3)  $X$  ha la  $p$ -MCP per un qualche  $p \in (1, \infty)$ .

*Sketch.* <sup>8</sup> È sufficiente dimostrare che 1 implica 2 e 3 implica 1 (poiché ovviamente 2 implica 3). Per dimostrare che 1 implica 2 si procede come segue: data una martingala  $(X_n, \mathcal{F}_n)$  uniformemente integrabile e limitata in  $L^1$  definiamo la funzione insiemistica

<sup>7</sup>Il risultato risulta più chiaro una volta osservato che  $L^1(\Omega, X) = L^1(\Omega) \hat{\otimes}_\pi X$ . Per uno studio approfondito del prodotto tensoriale di spazi di Banach e la sua relazione con la RNP si veda [4]

<sup>8</sup>Per una dimostrazione completa consultare [2]

$\mu$  mediante

$$\mu(A) := \lim_{n \rightarrow \infty} \int_A X_n d\mathbb{P}.$$

Si può dimostrare che tale definizione è ben posta e che  $\mu$  è una misura assolutamente continua rispetto a  $\mathbb{P}$ : dalla RNP segue che  $d\mu = f d\mathbb{P}$  ed è evidente che  $X_n$  converge in norma  $L^1$  e quasi certamente ad  $f$ . Dimostriamo ora che 3 implica 1 e siano dunque  $\mu$  una misura a valori vettoriali e  $\lambda$  una misura scalare definite su un comune spazio di misura  $(\Omega, \Sigma)$  tale che  $\mu \ll \lambda$ . Un semplice appello al teorema di Radon-Nykodim scalare ci permette di ridurci al caso  $\lambda = \|\mu\|$ , che possiamo supporre essere una probabilità senza perdita di generalità. Per ogni partizione finita  $\mathcal{B}$  di  $\Omega$  definiamo

$$f_{\mathcal{B}}(\omega) = \sum_k \frac{\mu(B_k)}{\|\mu\|(B_k)} 1_{B_k}(\omega).$$

Si dimostra facilmente che  $f_{\mathcal{B}}$  è una martingala limitata in  $L^\infty$  e quindi per (3) converge quasi certamente ed in norma ad  $f \in L^p$ . Ne segue la convergenza di  $f_{\mathcal{B}}$  ad  $f$  anche in  $L^1$  (in quanto  $L^p \hookrightarrow L^1$  continuamente). La continuità del valore atteso condizionato implica che per ogni  $A \in \Sigma$  si ha (denotiamo con  $\mathcal{C}$  la  $\sigma$ -algebra generata da  $A$ , i.e.  $\mathcal{C} = \{\emptyset, A, \Omega \setminus A, \Omega\}$ )

$$\mathbb{E}(1_A f) = \mathbb{E}(\mathbb{E}^{\mathcal{C}}(f)1_A) = \lim \mathbb{E}(\mathbb{E}^{\mathcal{C}}(f_{\mathcal{B}})1_A) = \mathbb{E}(f_{\mathcal{C}}1_A) = \mu(A)$$

□

PROPOSIZIONE 17.  $L^1(\{-1, 1\}^{\mathbb{N}}, \mathbb{P})^9$  non ha la RNP

*Dimostrazione.* Dimostriamo che non possiede la  $\infty$ -MCP. Per farlo, siano  $\pi_k \in L^1$  le funzioni di proiezione sui fattori ( $\pi_k((\omega_0, \omega_1, \dots, \omega_k, \dots)) = \omega_k$ ). Definiamo la martingala  $\{X_n\}$  associata alla filtrazione  $\{\mathcal{F}_n\}$ ;  $\mathcal{F}_n := \sigma(\pi_0, \dots, \pi_n)$  mediante

$$X_n(\omega) = \prod_{k \leq n} (1 + \pi_k(\omega)\pi_k).$$

Non è difficile verificare che si tratta di una martingala a valori in  $L^1$ . Tuttavia sfruttando l'indipendenza delle  $\pi_k$  otteniamo

$$\begin{aligned} \|X_{n+1}(\omega) - X_n(\omega)\| &= \mathbb{E} \left[ \prod_{k \leq n} (1 + \pi_k(\omega)\pi_k(\omega')) |\pi_{n+1}(\omega)| |\pi_{n+1}(\omega')| \right] \\ &= \mathbb{E} \left[ \prod_{k \leq n} (1 + \pi_k(\omega)\pi_k(\omega')) \right] \\ &= \prod_{k \leq n} \mathbb{E} [1 + \pi_k(\omega)\pi_k(\omega')] \\ &= 1 \end{aligned}$$

Di conseguenza la  $\infty$ -MCP fallisce. □

<sup>9</sup> $\mathbb{P}$  è la probabilità prodotto indotta munendo i fattori dell'usuale probabilità uniforme



OSSERVAZIONE 18. Come vedremo a breve,  $\ell_1$  ha la RNP in quanto duale separabile di  $c_0$ , quindi alcuni spazi  $L^1(\mu)$  hanno la RNP. D'altro canto, una vasta famiglia di spazi  $L^1(\mu)$  è isomorfa a  $L^1(\{-1, 1\}^{\mathbb{N}})$ : tutti gli spazi tali per cui  $\mu$  sia finita, separabile e priva di atomi<sup>10</sup>. In particolare,  $L^1([0, 1])$  (dov'è sottintesa la misura di Lebesgue) non ha la RNP.

COROLLARIO 19. *La RNP è separabilmente determinata:  $X$  ha la RNP sse la ha ogni suo sottospazio separabile*

*Dimostrazione.* Dal criterio di Pettis, la  $\infty - MCP$  è separabilmente determinata.  $\square$

TEOREMA 20. *Sia  $X$  uno spazio di Banach tale che  $X'$  è separabile.  $X'$  ha la RNP. Se  $X$  è riflessivo, ha la RNP.*

*Dimostrazione.* Dimostriamo che  $X'$  ha la  $\infty - MCP$  che per il teorema precedente è equivalente alla RNP. Sia dunque  $X_n$  una martingala limitata a valori in  $X'$ . Per il teorema di Banach-Alaogou esiste, per ogni  $\omega$ , un elemento  $f(\omega)$  di  $X'$  che è punto di accumulazione di  $X_n(\omega)$ . Sia ora  $E = \{x_1, \dots, x_n\}$  un sottoinsieme numerabile denso di  $X$ : per l'usuale teorema di convergenza di Doob abbiamo che per ogni  $k$ ,  $X_n(\omega)(x_k)$  converge quasi certamente a  $f(\omega)(x)$ . Poiché  $E$  è numerabile, abbiamo che per quasi ogni  $\omega \forall k X_n(\omega)(x_k) \rightarrow f(\omega)(x)$  e quindi  $X_n \xrightarrow{\sigma(X', X)} f$ . Rimane da dimostrare la forte misurabilità di  $f$ . Per dimostrarlo osserviamo innanzitutto che per ogni  $y \in X'$  la funzione  $\|f(\omega) - y\| = \sup_{x \in E: \|x\| \leq 1} f(\omega)(x) - y(x)$  è misurabile in quanto sup di una famiglia numerabile di funzioni misurabili; da questo che  $f^{-1}(B)$ , dove  $B$  è una bolla, è misurabile; poiché  $X'$  è separabile ne segue che  $f$  è misurabile (e dal criterio di Pettis sappiamo che è anche fortemente misurabile). Per dimostrare il secondo asserto osserviamo che discende dal primo e dal fatto che la RNP è separabilmente determinata.  $\square$

#### 4. Il teorema di Choquet alla Edgar

TEOREMA 21 (Edgar (3)). *Sia  $X$  uno spazio di Banach separabile dotato della RNP e sia  $C$  un sottoinsieme convesso chiuso e limitato di  $X$ . Dato  $x_0 \in C$  esiste  $\mu_{x_0}$  una misura con segno, finita e di Radon supportata su  $\text{ext}(C)$  tale che per ogni funzionale lineare  $f \in X'$  abbiamo*

$$f(x_0) = \int_{\text{ext}(C)} f d\mu_{x_0} \tag{2}$$

*Dimostrazione.* La separabilità è necessaria per poter utilizzare il teorema di selezione misurabile di Von Neumann. Più precisamente, osserviamo che  $C \times C \setminus \Delta$  è uno spazio polacco e dunque la mappa continua  $h : C \times C \setminus \Delta \rightarrow C \setminus \text{ext}(C)$  definita mediante  $h(x, y) = \frac{x+y}{2}$  ammette una selezione universalmente misurabile<sup>11</sup>  $g : C \setminus \text{ext}(C) \rightarrow$

<sup>10</sup>Il risultato è dovuto a Caratheorody in questa forma ed è noto come "teorema di isomorfismo". Esso è anche un caso particolare del teorema di Maharam, risultato molto più profondo nella teoria delle measure algebras

<sup>11</sup>ricordiamo che universalmente misurabile vuol dire che è misurabile nel senso di Carathoeodory per ogni misura boreliana finita

$C \times C \setminus \Delta$ . Estendendo  $g$  a  $\Delta$  mediante  $g(x, x) = x$ , otteniamo  $g : C \rightarrow C^2$ ; siano  $g_0$  e  $g_1$  le due componenti della funzione. L'idea (a dir poco geniale) di Edgar è la seguente: definiamo per induzione transfinita su  $\omega_1$  la seguente martingala su  $\Omega = \{0, 1\}^{\omega_1}$  (denotiamo con  $\pi_\alpha$  la proiezione sulla coordinata  $\alpha$ -esima). La costruzione avverrà come segue: per induzione transfinita su  $\alpha$ , costruiremo  $X_\alpha$  e dimostreremo che per ogni  $\gamma < \alpha$  si ha

$$\mathbb{E}(X_\alpha | \mathcal{F}_\gamma) = X_\gamma. \quad (3)$$

- (1)  $X_0 = x_0$ ,  $\mathcal{F}_0 = \{\emptyset, \Omega\}$

Che soddisfi (3) è ovvio

- (2) Se  $\alpha = \beta + 1$ ,  $X_\alpha = g_{\pi_\beta(\omega)}(X_\beta)$ ,  $\mathcal{F}_\alpha = \sigma(\pi_\lambda; 0 \leq \lambda < \alpha)$

Per dimostrare che soddisfa (3) cominciamo con  $\gamma = \beta$  nell'asserto. In tal caso l'asserto si dimostra osservando che preso  $A \in \mathcal{F}_\beta$

$$\begin{aligned} \mathbb{E}(1_A X_\alpha) &= \frac{1}{2} \mathbb{E}(1_A X_\alpha 1_{\pi_\beta=0}) + \frac{1}{2} \mathbb{E}(1_A X_\alpha 1_{\pi_\beta=1}) \\ &= \mathbb{E} \left( 1_A \frac{g_0(X_\beta) + g_1(X_\beta)}{2} \right) \\ &= \mathbb{E}(1_A X_\beta). \end{aligned}$$

Per gli altri valori di  $\gamma$  il risultato segue osservando che

$$\mathbb{E}(X_\alpha | \mathcal{F}_\delta) = \mathbb{E}(\mathbb{E}(X_\alpha | \mathcal{F}_\beta) | \mathcal{F}_\delta) = \mathbb{E}(X_\beta | \mathcal{F}_\delta)$$

- (3) Se  $\alpha$  è un ordinale limite,  $X_\alpha = \lim_{\lambda < \alpha} X_\lambda$  (il limite esiste grazie al teorema di Doob),  $\mathcal{F}_\alpha = \sigma(\pi_\lambda; 0 \leq \lambda < \alpha)$ .

Che valga (3) segue dal teorema di convergenza dominata:

$$\mathbb{E}(X_\alpha | \mathcal{F}_\gamma) = \lim_{\beta \rightarrow \alpha} \mathbb{E}(X_\beta | \mathcal{F}_\gamma) = X_\gamma$$

Di conseguenza,  $\{X_\alpha\}_{\alpha < \omega_1}$  è una martingala su  $\{0, 1\}^{\omega_1}$ . Costruiamo ora una funzione  $\psi$  strettamente convessa su  $C$ : poiché  $X$  è separabile, esiste un sottoinsieme  $E = \{f_1, \dots, f_n, \dots\}$  denso in  $(B_{X'}, \sigma(X', X))$  ed è sufficiente prendere

$$\psi(x) := \sum_{k=1}^{\infty} \frac{|f_k(x)|^2}{2^k}.$$

Costruiamo ora  $\Psi := \frac{1}{2} (\psi(g_0(x)) + \psi(g_1(x)))$ ; è evidente che  $\Psi \geq \psi$  con uguaglianza sse  $x \in \text{ext}(C)$  per stretta convessità; il nostro obiettivo è dimostrare che  $X_\alpha$  ha supporto contenuto in  $\text{ext}(C)$  per  $\alpha$  sufficientemente grande, da cui seguirà l'asserto prendendo come misura la legge di  $X_\alpha$ . Per dimostrare tale concentrazione del supporto è sufficiente dimostrare che

$$\mathbb{E}(\Psi(X_\alpha)) = \mathbb{E}(\psi(X_\alpha)) =: y_\alpha$$

per  $\alpha$  suff. grande. È evidente dalla costruzione di  $\Psi$  che

$$y_\alpha \leq \mathbb{E}(\Psi(X_\alpha)) = y_{\alpha+1},$$

quindi  $y_\alpha$  è una successione crescente su  $\omega_1$  ed è quindi definitivamente costante. Il risultato è dimostrato.  $\square$

OSSERVAZIONE 22. Choquet-Edgar implica direttamente che  $L^1(\mu)$  con  $\mu$  diffusa e separabile non ha la RNP:  $\text{ext}(B_{L^1(\mu)}) = \emptyset$ .

Come osservato da Edgar stesso, il teorema precedente si basa sull'utilizzo di 3 strumenti: l'esistenza della sezione misurabile della mappa  $h$ , il teorema di convergenza di martingala e l'esistenza di una funzione strettamente convessa. Ma è facile vedere che questi strumenti sono disponibili anche nel caso originale del teorema di Choquet e abbiamo quindi:

TEOREMA 23 (Choquet). *Sia  $K$  un sottoinsieme compatto convesso metrizzabile di un LC TVS  $V$ . Per ogni  $x \in K$  esiste una probabilità di Radon  $\mu_x$  supportata su  $\text{ext}(K)$  tale che per ogni  $y \in V'$*

$$y(x) = \int_{\text{ext}(K)} y(t) d\mu_x(t) \tag{4}$$

### 5. Bonus: il moto browniano

In questa ultima sezione vediamo, molto brevemente, come applicare il moto browniano all'analisi. L'argomento è estremamente vasto nella sua interezza e certamente non gli è fatta giustizia con questa breve introduzione: il lettore interessato può consultare [2,5].

TEOREMA 24. *Sia  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  una funzione armonica. Essa soddisfa la proprietà del valor medio, ovvero:*

$$u(x_0) = \int_{\partial B(x_0,r)} u(x) d\mathcal{H}^{n-1}. \tag{5}$$

Inoltre

$$u(x_0) = \int_{B(x_0,r)} u(x) d\mathcal{H}^n. \tag{6}$$

*Dimostrazione.* È sufficiente dimostrare la prima equazione. Possiamo supporre  $x_0 = 0, r = 1$ . Sia  $W_t$  un moto Browniano  $n$ -dimensionale che parte da 0. Grazie alla formula di Ito, otteniamo (poiché  $\Delta u = 0$ ) che  $u(W_t)$  è una martingala locale. Sia  $\tau$  il first hitting time di  $\partial B(0, 1)$ : si vede facilmente che  $u(W_{t \wedge \tau})$  è una martingala. Di conseguenza

$$\mathbb{E}[u(W_\tau)] = \mathbb{E}[u(W_0)] = u(0).$$

L'invarianza rotazionale del moto Browniano implica che  $W_\tau$  deve essere uniformemente distribuito su  $\partial B(0, 1)$  e l'asserto segue. □

OSSERVAZIONE 25. Il risultato è un corollario del teorema di Kakutani che collega misura armonica e moto Browniano.

Dalla proprietà del valor medio si ottiene facilmente il teorema di Liouville: è sufficiente osservare che  $|u(x) - u(y)| \lesssim \|u\|_\infty \frac{\mathcal{H}^n(B(x,r) \cap B(y,r))}{r^n} \rightarrow 0$ . È cionondimeno interessante osservare l'esistenza di una dimostrazione puramente probabilistica del risultato in due dimensioni:

TEOREMA 26 (Liouville). Sia  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$  una funzione armonica limitata. Essa è costante.

*Dimostrazione.* Seguendo lo stesso ragionamento della dimostrazione precedente otteniamo che  $u(W_t)$  è una martingala limitata. Per il teorema di convergenza di Doob abbiamo quindi che  $u(W_t)$  converge quasi certamente; poiché il moto Browniano in due dimensioni è ricorrente  $u$  deve essere costante.  $\square$

È possibile inoltre dimostrare il piccolo teorema di Picard utilizzando il moto Browniano. Della dimostrazione è fornito solamente uno sketch, il lettore interessato può consultare [5].

TEOREMA 27. Sia  $f : \mathbb{C} \rightarrow \mathbb{C} \setminus \{-1, 1\}$  intera. Essa è costante.

*Sketch.* Supponiamo wlog che  $f(0) = 0$  e sia  $W_t$  un moto browniano. Si supponga per assurdo  $f$  non costante. Il teorema di caratterizzazione del moto Browniano di Levy implica che  $f(W_t)$  è, a meno di cambio di parametrizzazione temporale, un moto Browniano. L'idea è la seguente: se  $f$  è non costante,  $f(\gamma)$ , dove  $\gamma$  è una qualunque curva chiusa, deve essere omotopa ad un punto; sfruttando le proprietà di ricorrenza e il teorema del supporto del moto Browniano è possibile costruire  $\gamma$  in maniera tale che  $f(\gamma)$  non sia omotopa ad un punto ottenendo una contraddizione.  $\square$

OSSERVAZIONE 28. La restrizione a una dimensione complessa non è dovuta ad una limitazione degli strumenti che stiamo utilizzando: esistono domini limitati in  $\mathbb{C}^2$  biolomorfi a  $\mathbb{C}^2$  (essi sono noti come domini di Fatou-Bieberbach).

## 6. Conclusione

Alcuni approfondimenti per il lettore interessato:

- (1) La teoria delle martingale (continue) può essere applicata molto proficuamente all'analisi armonica e complessa, in particolare agli spazi  $H^p$ . Si arriva addirittura a dimostrare il teorema della Corona così. Vedasi [5].
- (2) L'estensione di spazi di funzioni classici a funzioni a valori vettoriali non termina con  $L^p$ <sup>12</sup>. L'argomento è molto vasto e unisce tecniche e questioni di analisi armonica con problemi di analisi funzionale in maniera molto interessante, vedasi [6].
- (3) Degna di nota è la teoria degli spazi di Hardy a valori in spazi di Banach: qui RNP perde il suo ruolo centrale, sostituita dalla ARNP (analytic Radon-Nikodym property); si veda [2] per dettagli.
- (4) La teoria delle misure a valori vettoriali è incredibilmente ricca e va ben oltre i pochissimi risultati analizzati in queste pagine. Un ottimo punto di partenza per lo studio di questo argomento è la splendida monografia [1].

In chiusura, un sentito ringraziamento a tutti coloro che hanno reso possibile Fuori Orario in questi 3 anni durante i quali ho avuto il piacere di partecipare.

---

<sup>12</sup>come sa il lettore interessato alla teoria delle PDE, gli spazi di Sobolev a valori in spazi di Banach sono un utile strumento

**Bibliografia**

- [D] J. DIESTEL, J.J. UHL, *Vector Measures*
- [P] G. PISIER, *Martingales in Banach Spaces*
- [E] G.A. EDGAR, *A noncompact Choquet theorem*
- [R] R.R. RYAN, *Introduction to Tensor products of Banach spaces*
- [B] F. BASS, *Probabilistic techniques in analysis*
- [H] T. HYTÖNEN, J. VAN NEERVEN, M. VERAAR, L. WEIS, *Analysis in Banach Spaces*

Fuori Orario è una giornata di seminari *per studenti da studenti*, giunta quest'anno alla sua sesta edizione, organizzata dai giovani del Dipartimento di Matematica dell'Università degli Studi di Milano. Nata per condividere specifici argomenti di notevole bellezza che non trovano spazio all'interno delle lezioni ordinarie, questa iniziativa permette a noi studenti di via Saldini di metterci in gioco. Dopo due anni siamo finalmente tornati a svolgerlo in presenza nelle aule dell'università. Siamo così stati in grado di condividere le nostre idee e confrontarci dal vivo con grande passione.

*Stampato con il contributo dell'Università derivante dai fondi previsti per le attività culturali e sociali.*