

Frattali autosimili

1. Introduzione

Queste pagine riguardano una parte di un breve corso di approfondimento tenuto a una classe di studenti medi e mostrano una applicazione del ben noto teorema di Banach sulle contrazioni, trattato in una lezione precedente. Ne diamo qui un breve richiamo ricordando innanzi tutto che, in generale, se S è un insieme non vuoto e f è una applicazione di S in sé, le soluzioni dell'equazione

$$(1.1) \quad f(x) = x$$

si chiamano punti fissi per f .

Definizione 1.1. *Siano (S, d) uno spazio metrico completo e f una applicazione di S in sé. Diciamo che f è una contrazione in S se esiste un numero reale $\alpha < 1$ tale che*

$$(1.2) \quad d(f(x), f(y)) \leq \alpha d(x, y) \quad \forall x, y \in S.$$

Teorema 1.2 (delle contrazioni). *Siano (S, d) uno spazio metrico completo e f una contrazione in S . Allora f ha uno e un solo punto fisso. Inoltre, scelto $x_0 \in S$ ad arbitrio e definiti ricorsivamente x_1, x_2, \dots mediante la formula*

$$(1.3) \quad x_{n+1} = f(x_n), \quad n = 0, 1, 2, \dots$$

la successione $n \mapsto x_n$ converge al punto fisso. ■

L'applicazione del teorema riguarda la costruzione di frattali autosimili. Questi insiemi strani intervengono in modo naturale in varie questioni. Ecco un esempio tipico.

Consideriamo una funzione f di \mathbb{R} in \mathbb{R} e, a partire da un punto x_0 , definiamo successivamente x_1, x_2, \dots mediante la formula ricorrente

$$(1.4) \quad x_{n+1} = f(x_n)$$

Se f è una contrazione, la successione converge per il Teorema delle contrazioni, dato che la retta è uno spazio metrico completo rispetto alla distanza usuale. Se invece f non è una contrazione, il comportamento della successione dipende in generale dalla scelta di x_0 e si può incappare in situazioni difficili da trattare anche in casi apparentemente innocui, come quello della funzione data dalla formula

$$f(x) = 4x(1 - x), \quad x \in \mathbb{R}$$

che si annulla nei punti $x = 0, 1$, è positiva per $0 < x < 1$ e negativa altrove.

Dello stesso tipo, ma più abbordabile, è il caso della funzione “rettilenea a tratti”

$$(1.5) \quad g(x) = \begin{cases} 3x & \text{se } x \leq 1/2 \\ 3(1 - x) & \text{se } x > 1/2. \end{cases}$$

Si può fare qualche esperimento. Nella tabella che segue sono riportati i primi elementi x_0, x_1, x_2, \dots delle successioni costruite a partire da varie scelte di x_0

x_0	x_1	x_2	x_3	x_4	x_5	x_6	...
2	-3	-3 ²	-3 ³	-3 ⁴	-3 ⁵	-3 ⁶	...
1/2	3/2	-3/2	-3 ² /2	-3 ³ /2	-3 ⁴ /2	-3 ⁵ /2	...
1/4	3/4	3/4	3/4	3/4	3/4	3/4	...
1/13	3/13	9/13	12/13	3/13	9/13	12/13	...
1/40	3/40	9/40	27/40	39/40	3/40	9/40	...

ed è naturale chiedersi che cosa abbiano in comune, ad esempio, 1/4, 1/13 e 1/40 in riferimento alla funzione (1.5) considerata. Ne riparliamo. Altre scelte di x_0 forniscono poi successioni i cui elementi sembrano estratti a sorte ma, sorprendentemente, appartengono tutti all'intervallo $[0, 1]$. ■

In generale lo studio di successioni definite per ricorrenza mediante formule del tipo della (1.4) porta a considerare insiemi strani, finemente bucherellati o addirittura sbriciolati, se vogliamo usare due termini suggestivi. Questi insiemi, che intervengono anche in problemi di tipo completamente diverso, sono detti *frattali* e alcuni di essi hanno la proprietà di essere *autosimili*, cioè di essere costituiti da copie rimpicciolite di loro stessi. L'analisi sconfinava dunque nella geometria. Per questo motivo, ritenendo che al lettore sia più familiare considerare trasformazioni del piano anziché della retta, preferiamo trattare con qualche dettaglio il caso in cui l'ambiente sia il piano.

2. Frattali autosimili

Nel seguito il simbolo d denota la distanza usuale nel piano e la nozione di convergenza di successioni di punti del piano è intesa rispetto alla distanza d . Inoltre, per evitare ripetizioni frequenti, sottintendiamo che tutti gli insiemi che intervengono siano non vuoti.

Definizione 2.1. Sia λ un numero reale positivo. Una similitudine di rapporto λ è una funzione f del piano in sé che verifica la condizione

$$(2.1) \quad d(f(x), f(y)) = \lambda d(x, y)$$

per ogni coppia di punti x e y del piano. ■

Osservazione 2.2. Casi particolari di similitudini sono le *omotetie*. L'omotetia (più propriamente l'omotetia diretta) di centro O e rapporto λ è la trasformazione f costruita come segue. Si pone $f(O) = O$ e, per $x \neq O$, si definisce $f(x)$ come l'unico punto della semiretta uscente da O e contenente x che dista $\lambda d(x, O)$ da O . ■

A partire da N similitudini f_1, \dots, f_N del piano costruiamo una funzione che opera non tanto sui punti del piano quanto piuttosto sui suoi sottoinsiemi.

Definizione 2.3. Siano date N similitudini f_1, \dots, f_N del piano. Se A è un sottoinsieme del piano, denotiamo con $f(A)$ il sottoinsieme del piano descritto dalla condizione

seguinte: un punto y appartiene a $f(A)$ se e solo se esiste un punto x del piano la cui immagine tramite almeno una delle similitudini f_1, \dots, f_N sia y . ■

Dunque $f(A)$ si ottiene riunendo tutti i trasformati dell'insieme A tramite le N similitudini f_1, \dots, f_N assegnate.

Ci poniamo ora il problema di trovare gli insiemi che siano “punti fissi” per f , cioè gli insiemi I , detti anche *invarianti rispetto a f* (o, anche, invarianti rispetto alla data famiglia di similitudini), che verificano $f(I) = I$. Questo problema non può essere risolto in modo soddisfacente senza ipotesi ulteriori sulle similitudini di partenza e senza richieste ulteriori sull'insieme cercato. Basta infatti considerare il caso in cui $N = 1$. Se l'unica similitudine in gioco è una traslazione, per cui, in particolare, è $\lambda = 1$, allora ogni retta parallela alla direzione della traslazione è invariante rispetto a f , ma non è possibile trovare un insieme invariante che sia contenuto in un disco. Se la similitudine è invece, ad esempio, una rotazione di un quarto di giro intorno a un punto O assegnato, per cui ancora è $\lambda = 1$, addirittura ogni disco di centro O è un insieme invariante.

Per escludere i casi precedenti imporrò alle similitudini di avere rapporti < 1 e cercheremo i sottoinsiemi invarianti di tipo particolare, richiedendo che essi siano chiusi e limitati nel senso della definizione data di seguito.

Definizione 2.4. *Un sottoinsieme A del piano è detto chiuso quando verifica la condizione seguente: se $x \mapsto x_n$ è una qualunque successione di punti di A convergente, anche il limite della successione appartiene ad A . Un sottoinsieme del piano è detto invece limitato quando esiste un disco che lo contiene. ■*

Un disco, circonferenza che lo delimita compresa, è un insieme chiuso e limitato. Un esempio di insieme chiuso ma non limitato è dato da un semipiano, retta che lo individua compresa. Un esempio di insieme A limitato ma non chiuso è invece un disco, tolta la circonferenza che lo delimita: esistono infatti successioni di punti di A che convergono a punti che appartengono alla circonferenza e, dunque, non appartengono ad A .

Gli insiemi chiusi e limitati sono una buona versione piana della nozione di intervallo chiuso e per essi valgono i risultati che enunciamo di seguito.

Teorema 2.5 (di Weierstrass). *Siano A un insieme chiuso e limitato del piano e g una funzione di A in \mathbb{R} continua. Allora esistono almeno un punto x' e almeno un punto x'' di A tali che*

$$g(x') \leq g(x) \leq g(x'') \quad \forall x \in A. \quad \blacksquare$$

Naturalmente tali punti x' e x'' , che possono non essere unici, vengono detti *punti di minimo* e, rispettivamente, *punti di massimo* per g e per le loro immagini (valori minimo e massimo) vengono utilizzate le notazioni

$$\min_{x \in A} g(x) = g(x') \quad \text{e} \quad \max_{x \in A} g(x) = g(x'').$$

Teorema 2.6. *Siano A un insieme chiuso e limitato del piano e g una trasformazione continua del piano in sé. Allora anche l'immagine di A è un sottoinsieme chiuso e limitato del piano.*

Teorema 2.7. *Nelle condizioni e con le notazioni della Definizione 2.3, si supponga che le similitudini f_1, \dots, f_N abbiano rapporti $\lambda_1, \dots, \lambda_N < 1$. Allora esiste uno e un solo sottoinsieme chiuso e limitato I che sia invariante rispetto a f , cioè tale che $f(I) = I$. ■*

Traccia della dimostrazione. Utilizziamo il Teorema delle contrazioni di Banach. Come spazio metrico prendiamo naturalmente l'insieme S avente come elementi i sottoinsiemi chiusi e limitati del piano, munito di una distanza che ora definiamo. Come funzione alla quale applicare il teorema prendiamo la f stessa: gli insiemi invarianti rispetto a f sono infatti proprio i "punti" di S fissi per f . Per il Teorema 2.6, se A è chiuso e limitato, anche $f(A)$ lo è, per cui f risulta una applicazione di S in S .

La costruzione della distanza in S che rende applicabile il Teorema delle contrazioni è invece laboriosa. Fissiamo un punto y del piano e un elemento $A \in S$ e consideriamo la funzione

$$x \mapsto d(x, y), \quad x \in A.$$

Essa è continua in quanto

$$|d(x, y) - d(x', y)| \leq d(x, x') \quad \forall x, x'.$$

Applicandole allora il Teorema 2.5 di Weierstrass, deduciamo che essa ha almeno un punto di minimo. Esiste dunque il più piccolo dei valori $d(x, y)$ fra quelli ottenuti lasciando variare x in A . Poniamo allora

$$\text{dist}(y, A) = \min_{x \in A} d(x, y).$$

Il numero reale $\text{dist}(y, A)$ viene detto *distanza di y da A* . Al lettore è familiare il concetto analogo di distanza di un punto da una retta: qui la distanza è presa non da una retta ma da un insieme che appartiene a S , cioè da un insieme chiuso e limitato.

Ferma restando la scelta dell'insieme $A \in S$, lasciamo variare y nel piano e consideriamo la funzione

$$y \mapsto \text{dist}(y, A).$$

Ebbene, si dimostra che questa funzione è continua. In particolare, se fissiamo un altro elemento $B \in S$, cioè ancora un sottoinsieme B chiuso e limitato del piano, possiamo applicare il Teorema 2.5 di Weierstrass e dedurre che esiste il più grande dei valori $\text{dist}(y, A)$ fra quelli ottenuti lasciando variare y in B , cioè il valore

$$\max_{y \in B} \text{dist}(y, A).$$

In tutto il discorso fatto A e B sono due elementi arbitrari di S . In particolare i loro ruoli possono essere scambiati e sono ben definiti entrambi i numeri reali

$$(2.2) \quad \max_{y \in B} \text{dist}(y, A) \quad \text{e} \quad \max_{y \in A} \text{dist}(y, B).$$

A questo punto, finalmente, possiamo definire la distanza fra A e B come il più grande dei due numeri (2.2). Poniamo cioè

$$(2.3) \quad d_{\mathcal{H}}(A, B) = \max \left\{ \max_{y \in B} \text{dist}(y, A), \max_{y \in A} \text{dist}(y, B) \right\}.$$

La funzione $d_{\mathcal{H}}$ ora introdotta è effettivamente una distanza in S ed è detta *distanza di Hausdorff*. La nonnegatività e la simmetria sono evidenti, mentre la disuguaglianza triangolare è piuttosto laboriosa e complicata e non la controlliamo. Verifichiamo invece la proprietà rimanente, cioè che solo elementi uguali possono avere distanza nulla. Siano dunque $A, B \in S$ tali che $d_{\mathcal{H}}(A, B) = 0$: dobbiamo dimostrare che $A = B$. Siccome $d_{\mathcal{H}}(A, B) = 0$, i due numeri (2.2) sono entrambi nulli. Ora, l'annullamento del primo di essi significa

$$\text{dist}(y, A) = 0 \quad \forall y \in B.$$

Ma la definizione di $\text{dist}(y, A)$ fornisce un punto $x \in A$ tale che $\text{dist}(y, A) = d(x, y)$. Dunque, per ogni $y \in B$, esiste $x \in A$ tale che $d(x, y) = 0$, cioè un punto $x \in A$ tale che $x = y$. Ciò significa che ogni punto di B appartiene anche ad A . Usando l'annullamento del secondo dei numeri (2.2) deduciamo allo stesso modo che ogni punto di A appartiene anche a B . Dunque i due insiemi A e B coincidono.

I passi successivi della dimostrazione sono molto più difficili da controllare, per cui diamo solo gli enunciati. Primo: lo spazio metrico $(S, d_{\mathcal{H}})$ è completo. Secondo: vale la disuguaglianza

$$d_{\mathcal{H}}(f(A), f(B)) \leq \lambda d_{\mathcal{H}}(A, B) \quad \forall A, B \in S$$

ove λ è il più grande dei rapporti $\lambda_1, \dots, \lambda_N$ di similitudine. Siccome tutti questi rapporti sono < 1 per ipotesi, abbiamo $\lambda < 1$ e possiamo effettivamente applicare il Teorema delle contrazioni ottenendo l'esistenza e l'unicità del "punto fisso", cioè dell'insieme chiuso e limitato I tale che $f(I) = I$. ■

In realtà il Teorema delle contrazioni dice di più: se partiamo da un elemento qualunque di S , cioè da un qualunque sottoinsieme A_0 chiuso e limitato del piano, e costruiamo la successione $n \mapsto A_n$ con la formula ricorrente

$$(2.4) \quad A_{n+1} = f(A_n), \quad n = 0, 1, 2, \dots$$

troviamo una successione che converge all'insieme invariante I rispetto alla distanza $d_{\mathcal{H}}$ di Hausdorff. La cosa è sorprendente dato che l'insieme invariante è, di solito, molto complicato. Questo stesso insieme I , dunque, è ciò che si ottiene al limite anche quando si parte da un insieme A_0 bellissimo, come un disco o un quadrato.

L'insieme invariante I è particolarmente singolare ed è chiamato *frattale autosimile* quando le similitudini f_1, \dots, f_N godono di una proprietà ulteriore, una versione semplificata della quale è la seguente, che chiamiamo *condizione del quadrato*: esiste un quadrato Q tale che i quadrati Q_1, \dots, Q_N che si ottengono come immagini di Q tramite le similitudini f_1, \dots, f_N rispettivamente sono tutti contenuti in Q e hanno in comune a due a due al più punti del bordo (perimetro).

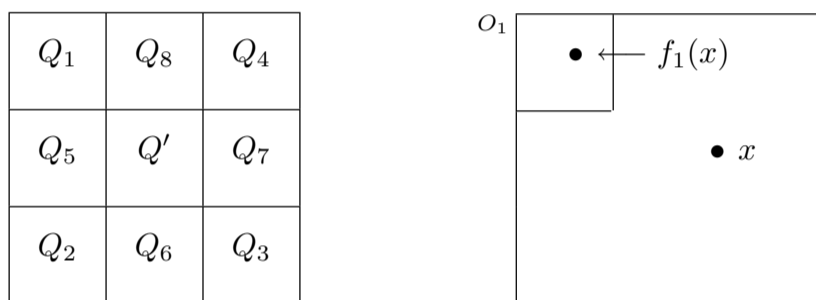
Se vale la condizione del quadrato l'insieme invariante I gode della proprietà seguente: le similitudini f_1, \dots, f_N trasformano I in copie rimpicciolite di I sostanzialmente prive di punti comuni e, dunque, I è un oggetto molto strano, del quale è difficile stabilire la dimensione, anche sul piano intuitivo. In molti casi, infatti, esso ha lunghezza infinita e area nulla (intuitivamente, ma la cosa potrebbe essere precisata) e i matematici gli

attribuiscono come dimensione un numero reale, di solito non intero. Questo è l'unica soluzione ρ dell'equazione

$$(2.5) \quad \lambda_1^\rho + \dots + \lambda_N^\rho = 1$$

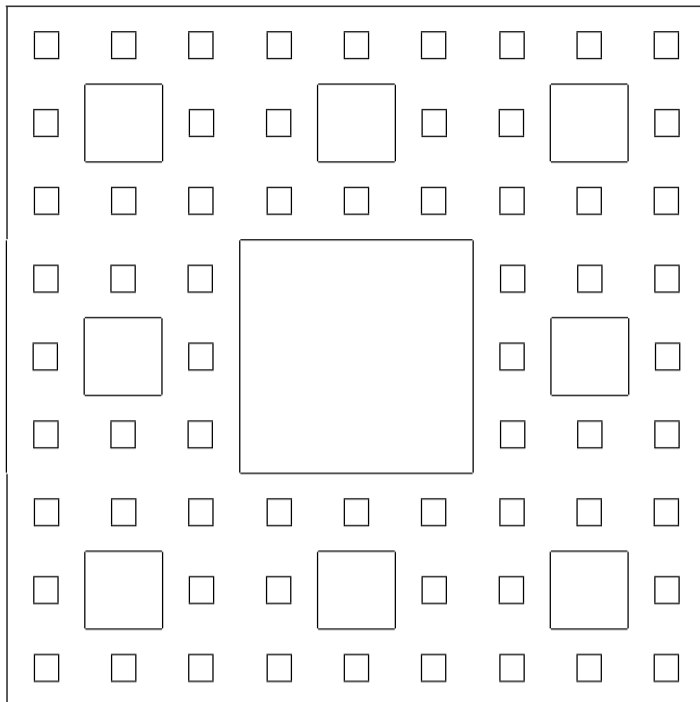
ed è chiamato *dimensione di similitudine di I* .

Esempio 2.8 (il tappeto di Sierpiński). Fissiamo un quadrato Q di lato ℓ e denotiamo con O_1, \dots, O_4 i suoi vertici e con O_5, \dots, O_8 i punti medi dei suoi lati (numerati in un ordine qualunque). Siano poi f_1, \dots, f_8 le omotetie di rapporto $1/3$ e di centri O_1, \dots, O_8 rispettivamente. Queste omotetie agiscono come segue. Suddividiamo Q in nove quadrati di lato $\ell/3$ mediante rette parallele ai lati di Q e, detto Q' il quadrato centrale, denotiamo con Q_1, \dots, Q_8 gli altri otto quadrati assumendo le notazioni in modo che il punto O_i appartenga a Q_i per $i = 1, \dots, 8$. Allora le omotetie f_1, \dots, f_8 trasformano Q in Q_1, \dots, Q_8 rispettivamente. Nelle due figure che seguono vengono descritte la suddivisione di Q nei nove quadrati e l'azione dell'omotetia f_1 .



Veniamo ora all'applicazione del Teorema 2.7, che è effettivamente applicabile in quanto i rapporti di similitudine valgono tutti $1/3$. Esiste dunque uno e un solo insieme chiuso e limitato I invariante rispetto alla trasformazione f costruita a partire dalle otto omotetie considerate. Esso è detto *tappeto di Sierpiński* e può essere ottenuto come limite, rispetto alla distanza di Hausdorff, di ciascuna delle successioni che si ottengono dalla (2.4) a partire da arbitrari insiemi A_0 chiusi e limitati.

Scegliamo ora come A_0 proprio il quadrato Q . La figura che segue rappresenta il terzo passo di questa costruzione. Il quadrato grande è Q e il quadrato centrale, di lato $\ell/3$, è quello che già abbiamo chiamato Q' . Si hanno poi 8 quadrati Q''_i , $i = 1, \dots, 8$, di lato $\ell/3^2$ e 8^2 quadrati Q'''_i , $i = 1, \dots, 8^2$, di lato $\ell/3^3$. La trasformazione f porta $A_0 = Q$ nell'insieme A_1 ottenuto riunendo i quadrati Q_1, \dots, Q_8 , cioè nell'insieme ottenuto togliendo Q' da Q . L'insieme successivo $A_2 = f(A_1)$ si ottiene allora da A_1 togliendo i quadrati Q''_i , $i = 1, \dots, 8$, e $A_3 = f(A_2)$ si ottiene togliendo da A_2 i quadrati Q'''_i , $i = 1, \dots, 8^2$. In definitiva A_3 si ottiene togliendo da Q tutti gli altri quadrati disegnati.



Se si immagina di proseguire nella costruzione e di passare al limite, si intuisce che il tappeto di Sierpiński è un insieme tutto bucherellato. La sua dimensione può essere calcolata applicando la (2.5) in quanto è soddisfatta la condizione del quadrato, come si vede prendendo proprio Q e gli otto quadrati Q_1, \dots, Q_8 nella verifica. Abbiamo

$$8 \cdot (1/3)^\rho = 1, \quad \text{cioè} \quad 3^\rho = 8, \quad \text{da cui} \quad \rho = \log_3 8 = 1.89\dots$$

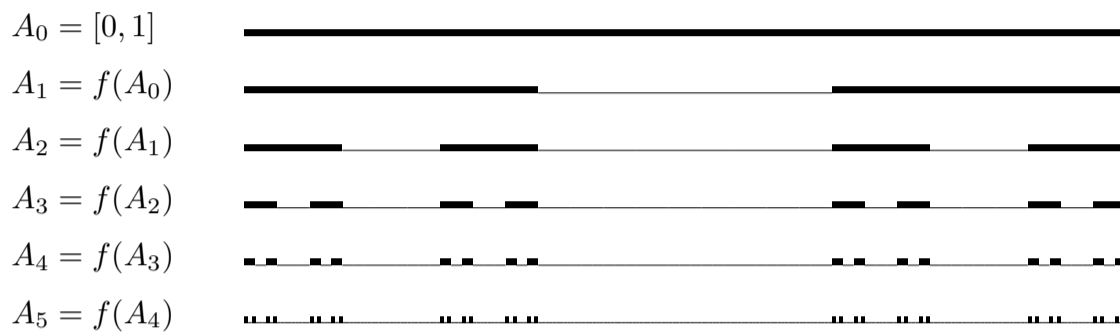
La stessa condizione del quadrato assicura poi che il tappeto è autosimile, cioè composto da otto parti, sostanzialmente prive di punti comuni e simili al tappeto intero.

Esempio 2.9 (l'insieme di Cantor). Di quanto abbiamo detto nel paragrafo si può dare la versione per i sottoinsiemi della retta anziché del piano. Consideriamo allora le due omotetie della retta di centri 0 e 1 e di rapporto $1/3$, le quali avvicinano il generico punto al loro centro riducendo tutte le distanze a un terzo del loro valore originario. Ad esempio, l'intervallo $[0, 1]$ viene trasformato nei due intervalli $[0, 1/3]$ e $[2/3, 1]$ rispettivamente.

L'analogo del Teorema 2.7 vale e la corrispondente trasformazione f ha uno e un solo insieme (chiuso e limitato) invariante, che denotiamo con \mathcal{C} e che è detto *insieme di Cantor*. Esso è, forse, il più semplice e famoso dei frattali autosimili.

Riprendendo le successioni costruite tramite la funzione (1.5), possiamo, se non chiarire il mistero, almeno portare un po' di luce. Si dimostra infatti che la funzione g data dalla (1.5) è in qualche modo legata alle due omotetie considerate (dunque alla funzione f da queste individuata) e che essa porta punti dell'insieme di Cantor ancora in punti dell'insieme di Cantor. Ebbene, i tre numeri $1/4$, $1/13$ e $1/40$, al contrario di quanto avviene per 2 e $1/2$, appartengono all'insieme di Cantor. Tuttavia la dimostrazione di queste affermazioni non è banale.

Veniamo ora alla costruzione dell'insieme di Cantor. Siccome esso è l'insieme invariante rispetto a f , l'insieme \mathcal{C} può essere ottenuto come limite di una successione di insiemi A_0, A_1, A_2, \dots costruita per mezzo della (2.4) a partire da un qualunque insieme chiuso e limitato A_0 della retta. Nella figura che segue sono riportati alcuni elementi della successione ottenuta scegliendo l'intervallo $[0, 1]$ come insieme A_0 iniziale.



Per quanto riguarda il calcolo della dimensione di \mathcal{C} , vale anche in questo caso la formula (2.5), purché sia soddisfatta la versione ad hoc della condizione del quadrato, versione che si ottiene sostituendo il termine “quadrato” con “segmento” e il termine “bordo” con “estremi”. Nel caso che stiamo esaminando possiamo prendere come segmento l'intervallo $[0, 1]$ e l'applicazione della (2.5) fornisce $\rho = \log_3 2 = 0.63 \dots$