

Matematiche complementari
A.A. 2021/2022

Ultimo aggiornamento:

22 giugno 2022

Indice

1	Prologo	2
2	Fondamenti teorici	3
2.1	Successioni di interi, sicurezza, strettezza, ...	5
2.2	Mappe tra spazi metrici	5
2.3	Similitudini	5
2.3.1	Forma canonica	6
2.4	Distanza di Hausdorff	7
2.5	Misure	8
2.6	Teorema fondamentale degli IFS	9
2.7	Misura di Hausdorff	11
2.8	Dimensione box	12
2.9	Proprietà di \dim_H e \dim_B	13
2.10	Tecniche per calcolare $\dim_H A$	16
2.11	Misure stazionarie	20
2.12	Dimensione di similitudine	23
2.13	Densità locale e similitudini	25
2.14	Barnsley-Vince: IFS topologici, etc.	28
3	Joint spectral radius e norme estreme	31
3.1	Raggio spettrale congiunto	31
3.2	Corpi convessi e norme	33
3.3	JSR e autoinsiemi	35
3.4	Norme di Protasov e Barabanov	36
4	Trasformazioni frattali	40
4.1	Frazioni continue e mappa di Farey	41
5	Gruppo di Möbius e geometria inversiva	41
5.1	Gruppi fuchsiani/kleiniani	48
5.2	Classificazione inversioni	49
5.3	Insiemi limite	50
5.4	Unificazione di cap e punti	50

1 Prologo

Definizione 1. Le *similitudini* di \mathbb{R}^n sono trasformazioni che hanno la forma

$$v \mapsto rOv + t, r \in \mathbb{R}_{>0}, t \in \mathbb{R}^n,$$

con $O \in O_n(\mathbb{R})$.

Osserviamo che stiamo usando l'intero gruppo ortogonale, per cui sono ammesse anche mappe che invertono l'orientazione dello spazio. Le similitudini formano un gruppo:

$$\text{Sim} = \text{Trasl} \rtimes \mathbb{R}_{>0} \times O_n \mathbb{R}.$$

Per capire cosa significa \rtimes , facciamo un piccolo ripasso di algebra. Una *sequenza esatta corta* è una successione di mappe tra gruppi

$$1 \rightarrow H \rightarrow G \rightarrow T \rightarrow 1$$

tali che l'immagine di ogni mappa è il nucleo della successiva. Per un noto teorema di isomorfismo, questo ci consente di dire che $T \simeq G/H$. Quindi a meno di isomorfismi, G è prodotto di H e T . Per mettere in evidenza che H è un sottogruppo normale, scriviamo quindi $H \rtimes T = G$.

Nel nostro caso particolare abbiamo la sequenza esatta

$$1 \rightarrow \text{Trasl} \rightarrow \text{Sim} \rightarrow \mathbb{R}_{>0} \times O_n \mathbb{R} \rightarrow 1$$

e si può vedere facilmente che $\text{Trasl} \triangleleft \text{Sim}$ perché se indichiamo come T_t la traslazione di t ,

$$GT_tG^{-1} = T_{G(t)} \quad \forall G \in \text{Sim}.$$

Infatti vale $G \circ T_t = T_{G(t)} \circ G$, e per vederlo basta considerare un vettore $v \in \mathbb{R}^n$ e ricordare che gli elementi di Sim sono lineari. Otteniamo

$$(G \circ T_t)(v) = G(v + t) = G(v) + G(t) = (T_{G(t)} \circ G)(v).$$

Possiamo rappresentare un generico elemento $v \mapsto rOv + t$ di Sim in forma matriciale come

$$\begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ 1 \end{pmatrix}$$

ponendo ovviamente $A := rO$. Questa rappresentazione risulta molto utile per implementare l'applicazione di similitudini in un calcolatore.

Definizione 2. Consideriamo un insieme $S := \{S_1, \dots, S_N\}$ di similitudini di \mathbb{R}^n . L'*operatore di Hutchinson associato a S* è la mappa da \mathbb{R}^n in sé definita da

$$\Phi_S(U) := S_1[U] \cup \dots \cup S_N[U]$$

dove $U \subseteq \mathbb{R}^n$.

Vediamo ora l'enunciato di un teorema fondamentale.

Teorema 1. Se tutte le similitudini S_i sono contrazioni (i.e., $0 < r_1, \dots, r_N < 1$) allora

- Φ_S un unico punto fisso, e questo è compatto;
- Partendo da un insieme compatto e iterando Φ_S , la successione di insiemi ottenuta converge (in una appropriata metrica) al punto fisso.

Osservazione 3. Se K è compatto, anche $\Phi_S(K)$ lo è. Infatti le similitudini sono continue, e l'unione finita di compatti è un compatto.

Esempio 1 (insieme di Cantor). Consideriamo \mathbb{R} e le due mappe

$$S_1 : x \mapsto \frac{1}{3}x \quad S_2 : x \mapsto \frac{1}{3}x + 2/3.$$

Ovviamente si tratta di contrazioni, per cui si applica il teorema enunciato prima. Posto $S = \{S_1, S_2\}$, vediamo che l'insieme di Cantor è un punto fisso per l'operatore di Hutchinson:

$$\Phi_S(C) = \Phi_S\left(\bigcap \{\text{livelli di } C\}\right) = \bigcap \Phi_S(\{\text{livelli di } C\}) = C.$$

2 Fondamenti teorici

L'ambiente entro il quale svilupperemo la teoria consta di uno spazio metrico completo (X, d) localmente compatto e *second countable*, ossia con base numerabile. Un esempio di questi spazi è il classico $(\mathbb{R}^d, \|\cdot\|_2)$. Spesso si dice che X è uno *spazio polacco* se è uno spazio topologico su cui si può mettere una metrica che lo rende completo (ad esempio $(0, 1)$ è polacco: è completo con certe metriche).

Definizione 4. Diremo che una funzione $F : X \rightarrow Y$ è *propria* se è continua e la controimmagine di ogni compatto è compatta.

Teorema 2. Sia $F : X \rightarrow Y$ propria e sia $(x_n) \subset X$ una successione la cui F -immagine converge a qualche $y \in Y$. Allora esiste una sottosuccessione (x_{n_k}) tale che $x_{n_k} \xrightarrow{k \rightarrow +\infty} x \in F^{-1}\{y\}$.

Dimostrazione. Poiché Y è localmente compatto, senza perdita di generalità possiamo supporre che tutta la successione (Fx_n) sia contenuta in un intorno compatto $K \ni y$. Dunque $(x_n) \subset F^{-1}[K]$ che è compatto perché F è propria, e dunque (per una nota caratterizzazione) è pure sequenzialmente compatto. Allora esiste una sottosuccessione per cui $x_{n_k} \rightarrow x \in F^{-1}[K]$. Poiché F è continua, abbiamo allora $Fx_{n_k} \rightarrow Fx$; d'altra parte, $Fx_{n_k} \rightarrow y$, per cui dall'unicità del limite segue che $Fx = y$. \square

Corollario 3. Ogni mappa propria è chiusa (cioè manda chiusi in chiusi).

Dimostrazione. Sia $C \subset X$ chiuso e sia y un punto di accumulazione per $F[C]$. Per definizione esiste una successione approssimante $Fx_n \rightarrow y$, che possiamo sollevare a X in modo che $x_n \in C, \forall n \in \mathbb{N}$. Per il teorema precedente, possiamo estrarre una sottosuccessione $x_{n_k} \rightarrow x \in F^{-1}\{y\}$. Essendo C chiuso si ha $x \in C$, ma allora $y = Fx \in F[C]$. \square

Corollario 4. Sia $F : X \rightarrow Y$ continua e iniettiva con X compatto. Allora $F|_X : X \rightarrow F[X]$ è un omeomorfismo.

Dimostrazione. Bisogna solo verificare che l'inversa è continua, o equivalentemente che F è aperta, o equivalentemente che F è chiusa. Questo è vero grazie al corollario precedente perché F è propria: infatti un compatto in Y è chiuso, per cui per continuità la sua controimmagine è chiusa in X e un chiuso in un compatto è compatto. \square

Definizione 5. Una contrazione $F : X \rightarrow X$ si dice *stretta* se $\exists \lambda \in [0, 1)$ t.c.

$$\forall x, x' \in X, d(Fx, Fx') \leq \lambda d(x, x').$$

Teorema 5 (delle contrazioni). Sia $F : X \rightarrow X$ una contrazione stretta sullo spazio metrico completo (X, d) . Allora F ha uno e un solo punto fisso, e questo è il limite della successione $x_0, x_1 = Fx_0, x_2 = Fx_1, \dots$

Dimostrazione. Consideriamo la successione definita nell'enunciato. Dato che F è una contrazione stretta, diciamo di costante λ , abbiamo che $d(x_2, x_1) = d(Fx_1, Fx_0) \leq \lambda d(x_1, x_0)$, e quindi per induzione

$$d(x_{n+1}, x_n) \leq \lambda^n d(x_1, x_0)$$

e il secondo membro tende a 0 per $n \rightarrow \infty$ dato che $\lambda \in [0, 1)$ e $d(x_1, x_0)$ è finita. Per usare la completezza dimostriamo che la successione (x_n) è di Cauchy. Siano quindi $m > n$ e consideriamo la seguente maggiorazione

$$\begin{aligned} d(x_m, x_n) &\leq \lambda^n d(x_{m-n}, x_0) \leq \\ &\leq \lambda^n (d(x_{m-n}, x_{m-n-1}) + \dots + d(x_1, x_0)) \leq \\ &\leq \lambda^n (\lambda^{m-n-1} + \dots + 1) d(x_1, x_0) \leq \\ &\leq \left(\sum_{k=0}^{+\infty} \lambda^k \right) \lambda^n d(x_1, x_0) = \frac{1}{1-\lambda} \lambda^n d(x_1, x_0) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

dove il limite segue per quello calcolato in precedenza e perché ovviamente $1/(1-\lambda)$ è finito. Quindi (x_n) è una successione di Cauchy in X che è uno spazio completo, per cui $x_n \xrightarrow{n \rightarrow \infty} \bar{x} \in X$.

Ora vediamo che tale limite è punto fisso per la contrazione F . Consideriamo infatti la relazione

$$d(\bar{x}, Fx_n) = d(\bar{x}, x_{n+1}).$$

Il termine di sinistra tende per continuità di F e d a $d(\bar{x}, F\bar{x})$, mentre il termine di destra tende (sempre per la continuità della distanza) a $d(\bar{x}, \bar{x}) = 0$. Per l'unicità del limite segue che $d(\bar{x}, F\bar{x}) = 0$.

Infine verifichiamo l'unicità: siano \bar{x}, \bar{x}' due punti fissi per F . Allora abbiamo, usando la definizione di contrazione,

$$d(\bar{x}, \bar{x}') = d(F\bar{x}, F\bar{x}') \leq \lambda d(\bar{x}, \bar{x}')$$

cioè $(1-\lambda)d(\bar{x}, \bar{x}') \leq 0$. Dato che $1-\lambda > 0$, $d(\bar{x}, \bar{x}') \leq 0$. Per la non negatività della distanza, questo implica che $d(\bar{x}, \bar{x}') = 0$, e quindi che $\bar{x} = \bar{x}'$. \square

Definizione 6. Uno spazio metrico si dice *localmente compatto* se ogni punto dello spazio ha un intorno la cui chiusura è compatta.

Esempio 2. Gli spazi \mathbb{R}^n e \mathbb{C}^n sono localmente compatti. Gli spazi di Hilbert su \mathbb{C} o su \mathbb{R} di dimensione infinita *non* sono localmente compatti (un esempio concreto è ℓ^2). Neanche \mathbb{Q} con la topologia indotta da \mathbb{R} è localmente compatto.

Per completezza (della spiegazione) ricordiamo alcuni risultati noti.

Teorema 6. Per uno spazio metrico (X, d) sono equivalenti:

- (X, d) è compatto;
- (X, d) è numerabilmente compatto;
- (X, d) è sequenzialmente compatto;

Teorema 7. Se $A \subseteq (X, d)$, A è compatto se e solo se A è chiuso e totalmente limitato. Se vale una, e quindi entrambe, le condizioni, A è chiuso e limitato.

Teorema 8. Uno spazio vettoriale topologico è localmente compatto se e solo se ha dimensione algebrica finita.

2.1 Successioni di interi, sicurezza, strettezza, ...

2.2 Mappe tra spazi metrici

Consideriamo la mappa $F : X \rightarrow X$, e definiamo la *costante di Lipschitz di F* come

$$\text{Lip}F = \sup_{x \neq y} \frac{d(F(x), F(y))}{d(x, y)}.$$

Segue dalla definizione che se $\text{Lip}F = \lambda$, allora $d(F(x), F(y)) \leq \lambda d(x, y)$ per ogni $x, y \in X$ e inoltre la costante di Lipschitz è la più piccola con tale proprietà. Diciamo che F è *lipschitziana* se $\text{Lip}F < \infty$ e che F è una *contrazione* se $\text{Lip}F < 1$. Come abbiamo dimostrato in precedenza, le contrazioni hanno un unico punto fisso.

Supponiamo di avere una famiglia finita $\mathcal{S} = \{S_1, \dots, S_N\}$ di mappe $S_i : X \rightarrow X$. Ci troveremo spesso a comporne una successione, e per rendere la notazione migliore scriveremo $S_{i_1 \dots i_p} := S_{i_1} \circ \dots \circ S_{i_p}$.

2.3 Similitudini

Diciamo che $S : X \rightarrow X$ è una *similitudine* se $d(S(x), S(y)) = rd(x, y)$ per ogni $x, y \in X$ per un certo r fissato. Focalizzandoci sul caso di mappe su \mathbb{R}^n , scriveremo $T_b : x \mapsto x + b$ per la *traslazione* e $D_r : x \mapsto rx$ per la *dilatazione* di fattore $r \geq 0$.

Proposizione 9. $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è una similitudine se e solo se si può scrivere come

$$S = D_r \circ T_b \circ O$$

per opportuni $r \geq 0$, $b \in \mathbb{R}^n$ e $O \in O_n \mathbb{R}$.

Dimostrazione. (\Leftarrow) è ovvio. (\Rightarrow) Sia quindi S una similitudine con $\text{Lip}S = r \neq 0$. Definiamo la mappa

$$g(x) = r^{-1}(S(x) - S(0)).$$

Essa è un'isometria che fissa l'origine $0 \in \mathbb{R}^n$. Applicando l'identità di polarizzazione, dato che preserva la norma essa preserva anche il prodotto scalare (*).

Sia ora $\{e_i\}$ una base ortonormale di \mathbb{R}^n . Per quanto appena osservato anche $\{g(e_i)\}$ è una base ortonormale. Quindi se prendiamo un generico $x = \sum \langle x, e_i \rangle e_i$, abbiamo che

$$g(x) = \sum \langle g(x), g(e_i) \rangle g(e_i) \stackrel{*}{=} \sum \langle x, e_i \rangle g(e_i)$$

cioè g è anche lineare. Questo, unito al fatto che sia un'isometria che fissa l'origine, implica che $g \in O_n \mathbb{R}$.

Ora è facile concludere perché invertendo la relazione che definisce g si ottiene

$$S(x) = r[g(x) + r^{-1}S(0)] = D_r \circ T_{r^{-1}S(0)} \circ g(x).$$

□

Osservazione 7. La stessa dimostrazione vale anche su spazi complessi, sostituendo le trasformazioni ortogonali con quelle unitarie. Infatti la cruciale identità di polarizzazione vale anche in questo caso.

2.3.1 Forma canonica

Proposizione 10. Consideriamo una similitudine $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ di ragione $0 < r < 1$, con punto fisso $a \in \mathbb{R}^n$. Posto $O^a := T_a \circ O \circ T_a^{-1}$ e $D_r^a := T_a \circ D_r \circ T_a^{-1}$ per una opportuna $O \in O_n \mathbb{R}$, vale

$$S = D_r^a \circ O_a.$$

Dimostrazione. Abbiamo $S = D_r^a \circ (D_r^a)^{-1} \circ S =: D_r^a \circ R$. Osserviamo che R fissa a ed è un'isometria. Quindi R è del tipo $T_a \circ O \circ T_a^{-1}$. □

Quindi ogni similitudine si può pensare come l'applicazione successiva di una trasformazione ortogonale centrata in a e di una dilatazione di ragione r sempre centrata in a . Quindi una similitudine si può caratterizzare dando tre elementi: trasformazione ortogonale, punto fisso e ragione. Questo porta a definire la *forma canonica*

$$S = (a, r, O).$$

Osservazione 8. Se consideriamo due similitudini $S_1 = (a_1, r_1, O_1)$ e $S_2 = (a_2, r_2, O_2)$, allora $S_1 \circ S_2 = (a, r, O)$ dove $r = r_1 r_2$, $O = O_1 \circ O_2$ e l'espressione per il nuovo punto fisso è

$$a = a_2 + (I - r_1 r_2 O_1 O_2)^{-1} (I - r_1 O_1) (a_1 - a_2).$$

2.4 Distanza di Hausdorff

Dato $x \in X$ e $A \subset X$, la *distanza* tra x e A è definita come

$$d(x, A) := \inf\{d(x, a) : a \in A\}.$$

Se $A \subset X$ e $\varepsilon > 0$, definiamo l' ε -*ingrandimento* di A come

$$A_\varepsilon := \{x \in X : d(x, A) < \varepsilon\}.$$

Definizione 9. La *distanza asimmetrica* $\vec{d}(A, B) = \sup\{d(a, B) : a \in A\} = \min\{\varepsilon > 0 : A \subseteq B_\varepsilon\}$, che si può pensare come il sup delle distanze che un ipotetico nemico mi obbliga a percorrere per andare in B partendo da A .

Simmetrizziamo la distanza, ottenendo la seguente definizione.

Definizione 10. La *distanza di Hausdorff* è

$$d(A, B) := \max\{\vec{d}(A, B), \vec{d}(B, A)\}.$$

Osservazione 11. In generale è solo una pseudometrica, infatti la non degenerazione della distanza fallisce considerando ad esempio $A = [0, 1]$ e $B = (0, 1]$. Tuttavia basta restringersi a considerare insiemi chiusi per ottenere una metrica valida (esercizio, l'unica cosa non ovvia è la disuguaglianza triangolare). Un'importante proprietà dello spazio dei chiusi limitati con la metrica di Hausdorff è che è completo. Inoltre se $K \subset X$ è compatto, anche $\{\text{compatti}\} \cap \{A : A \subset K\}$ è compatto.

Proprietà 11 (della metrica di Hausdorff). *Data* $F : X \rightarrow X$,

1. se $\text{Lip}F = r$, allora $\forall K_1, K_2$ compatti vale $d(FK_1, FK_2) \leq rd(K_1, K_2)$;
2. $d(\bigcup_{i \in I} A_i, \bigcup_{i \in I} B_i) \leq \sup_{i \in I} d(A_i, B_i)$.

Dimostrazione. Iniziamo con il primo punto, osservando che basta dimostrare la tesi per le distanze orientate e poi passando al massimo la tesi segue facilmente. Per prima cosa osserviamo che preso $x \in K_1$ vale $d(Fx, FK_2) \leq rd(x, K_2)$. Infatti dato che K_2 è compatto, esiste $y_0 \in K_2$ tale che $d(x, K_2) = d(x, y_0)$. Allora

$$d(Fx, FK_2) \leq d(Fx, Fy_0) \leq rd(x, y_0) = rd(x, K_2).$$

Quindi abbiamo che

$$\vec{d}(FK_1, FK_2) = \max_{x \in K_1} d(Fx, FK_2) \leq r \max_{x \in K_1} d(x, K_2) = r\vec{d}(K_1, K_2),$$

per cui si conclude come detto all'inizio.

Il secondo punto è facile, basta convincersi che nel l.h.s. si ha più scelta sui percorsi possibili sui cui minimizzare la distanza. \square

2.5 Misure

Nel corso di Probabilità I abbiamo visto la definizione di misura dovuta a Kolmogorov. Ora vediamo una generalizzazione data da Caratheodory.

Definizione 12. Una *misura esterna* μ^* su X è una mappa $\mu^* : \mathcal{P}(X) \rightarrow [0, +\infty]$ tale che

1. $\mu^*(\emptyset) = 0$;
2. $\mu^*\left(\bigcup_{i \in I} E_i\right) \leq \sum_{i \in I} \mu^*(E_i)$;
3. $A \subseteq B \implies \mu^*(A) \leq \mu^*(B)$.

Le ultime due condizioni possono essere equivalentemente sostituite da

$$2b. A \subseteq \bigcup_{i \in I} B_i \implies \mu^*(A) \leq \sum_{i \in I} \mu^*(B_i).$$

Osservazione 13. La misura esterna è definita su *tutti* gli elementi dell'insieme delle parti di X .

Come si collega tale definizione con quella di Kolmogorov? La risposta è nella seguente definizione.

Definizione 14. $A \subset X$ è *misurabile nel senso di Caratheodory* se $\forall T \subset X$ vale

$$\mu^*(T) = \mu^*(T \cap A) + \mu^*(T \cap A^C).$$

Si dimostra che tale classe di insiemi è una σ -algebra \mathcal{F} su X , e $\mu := \mu^*|_{\mathcal{F}}$ è una “vera” misura, cioè lo è nel senso di Kolmogorov.

Definizione 15. Una misura μ^* su X si dice *Borel-regolare* se tutti i boreliani di X sono misurabili e $\forall A \subset X$ esiste un boreliano $B \supset A$ tale che $\mu^*(A) = \mu^*(B)$.

In spazi metrici completi la regolarità si può esprimere in modo diverso (come fatto in Analisi II), potendo caratterizzare la misura di un insieme $A \subset X$ come estremo superiore sulle misure dei compatti contenuti in A o come estremo inferiore delle misure degli aperti che lo contengono.

Definizione 16. Denoteremo con \mathcal{M} l'insieme di misure Borel-regolari con supporto limitato e massa totale finita, e in particolare indicheremo le misure di tale insieme che sono delle probabilità con

$$\mathcal{M}^1 := \{\mu \in \mathcal{M} \mid \mathbf{M}(\mu) = 1\}.$$

Definizione 17. Definiamo $\mathcal{BC}(X) := \{f : X \rightarrow \mathbb{R} : f \text{ è continua e limitata sui limitati}\}$.

Considerando una misura $\mu \in \mathcal{M}$ e una $\phi \in \mathcal{BC}(X)$, definiamo $\mu(\phi) := \int \phi d\mu$. Si ha che $\mu : \mathcal{BC} \rightarrow [0, \infty)$, μ è lineare ed è positiva.

Se poi consideriamo una funzione $f : X \rightarrow X$ misurabile, possiamo definire il push-forward $f_* : \mathcal{M} \rightarrow \mathcal{M}$ come $(f_*\mu)(E) := \mu(f^{-1}[E])$. Equivalentemente, vale $(f_*\mu)(\phi) = \mu(\phi \circ f)$, ossia il teorema del portare in giro. Osserviamo che la massa totale si conserva, ossia $\mathbf{M}(f_*\mu) = \mathbf{M}(\mu)$.

Ora vediamo che si può immergere \mathcal{M} nell'insieme $\mathbb{R}^{\mathcal{BC}}$. Infatti se fissiamo un elemento $\mu \in \mathcal{M}$ abbiamo a disposizione per ogni $\phi \in \mathcal{BC}$ la proiezione sulla “ ϕ -esima” componente, $\pi_\phi(\mu) := \int \phi d\mu$. Resta quindi definita la mappa $\mu \mapsto (\pi_\phi(\mu))_\phi$ che associa ad una misura tutti i valori degli integrali delle funzioni in \mathcal{BC} fatte con essa. Tale associazione è iniettiva, e inoltre si vede che \mathcal{M} eredita la *topologia debole* da $\mathbb{R}^{\mathcal{BC}}$, che ha la topologia prodotto indotta da \mathbb{R} . Una base della topologia debole sono gli aperti $\{\mu : a < \mu(\phi) < b\}$ con $a < b$ reali e $\phi \in \mathcal{BC}$ arbitraria.

2.6 Teorema fondamentale degli IFS

Consideriamo un insieme finito $S = \{S_0, \dots, S_{N-1}\}$ di contrazioni con costanti di Lipschitz $0 < r_0, \dots, r_{N-1} < 1$. Definiamo poi $N^\omega = \{(i_0 i_1 \dots) =: \mathbf{i} \mid i_j \in \{0, \dots, N-1\}\}$ e useremo la seguente notazione $\mathbf{i} \upharpoonright t := i_0 \dots i_{t-1}$ per la restrizione ai primi t elementi della successione. Per $w \in N^{<\omega}$ finita, scriveremo $S_w := S_{i_0} \circ \dots \circ S_{i_{t-1}}$ e se $A \subseteq \mathbb{R}^n$, $A_w := S_w[A]$. Infine denotiamo con $\Phi : \mathcal{K} \rightarrow \mathcal{K}$ l'operatore di Hutchinson associato ad S .

Osservazione 18. Vale $\Phi^n(A) = \bigcup_{w \in N^n} A_w$, infatti

$$\begin{aligned} \Phi(A) &= S_0[A] \cup S_{N-1}[A] = A_0 \cup \dots \cup A_{N-1} \\ \Phi^2(A) &= S_0[S_0 A \cup \dots \cup S_{N-1} A] \cup \dots \cup S_{N-1}[S_0 A \cup \dots \cup S_{N-1} A] \\ &= \bigcup_{i,j} S_i S_j A = \bigcup_{w \in N^2} A_w \end{aligned}$$

e si conclude per induzione.

Teorema 12 (fondamentale di caratterizzazione degli IFS strettamente contrattivi). *Con le ipotesi fatte sopra, valgono i seguenti risultati.*

1. $\exists! K \in \mathcal{K}$ punto fisso per Φ .
2. $\forall w \in N^{<\omega}, K_w = K_{w0} \cup \dots \cup K_{w(N-1)}$.
3. $\forall \mathbf{i} \in N^\omega$, vale la catena

$$K = K_{\mathbf{i}|0} \supseteq K_{\mathbf{i}|1} \supseteq K_{\mathbf{i}|2} \supseteq \dots$$

e l'intersezione di tutti gli elementi della catena è un singoletto $\{\pi(\mathbf{i})\}$. La mappa $\pi : N^\omega \rightarrow K$ è suriettiva.

4. Per ogni $w \in N^{<\omega}$, sia s_w l'unico punto fisso di S_w . Allora $s_w = \pi(\bar{w})$, dove $\bar{w} = www \dots$. Inoltre $\pi(\mathbf{i}) = \lim_{n \rightarrow \infty} s_{\mathbf{i}|n}$.
5. Il punto fisso K è la chiusura topologica dei punti fissi delle composizioni finite, ossia $K = \{s_w : w \in N^{<\omega}\}^f$.
6. $S_w[A_u] = A_{wu}$ e $S_w(\pi(\mathbf{i})) = \pi(w\mathbf{i})$.
7. π è continua.

8. Sia $H \in \mathcal{K}$ qualsiasi. Allora $\Phi^n(H) \xrightarrow{n \rightarrow +\infty} K$ nella metrica di Hausdorff. Inoltre, con la distanza di Hausdorff, vale

$$\forall \mathbf{i} \in N^\omega, \quad d(\pi(\mathbf{i}), H_{\mathbf{i}|n}) \rightarrow 0$$

uniformemente, cioè con velocità indipendente da \mathbf{i} .

8'. La costruzione di $\pi(\mathbf{i})$ vale anche usando un qualsiasi $H \in \mathcal{K}$ tale per cui $\Phi(H) \subseteq H$.

Dimostrazione. Per (1) e la prima parte di (8) basta osservare che Φ è una contrazione stretta su \mathcal{K} e il risultato segue dal teorema delle contrazioni. In effetti presi $H, T \in \mathcal{K}$ si ha, usando le proprietà della distanza di Hausdorff, che vale

$$d(\Phi T, \Phi H) = d(\cup_i T_i, \cup_i H_i) \leq \max_i d(T_i, H_i) \leq \underbrace{\max_i r_i}_{=: r < 1} d(T, H) = r d(T, H).$$

Sia ora K il punto fisso di Φ . Per l'osservazione fatta prima dell'enunciato si ha $K = \bigcup_{u \in N^n} K_u$, e più in generale $K_w = \bigcup_{u \in N^n} K_{wu}$. Per dimostrare (3) va verificata solo la suriettività di π , in quanto gli altri risultati sono ovvi (basta osservare che il diametro degli insiemi della catena si restringe e ricordare che sono compatti). Sia dunque $x \in K = K_0 \cup \dots \cup K_{N-1}$. Dato che x sta nell'unione di tali insiemi, esiste (almeno) un $i_0 \in \{0, \dots, N-1\}$ per cui $x \in K_{i_0}$. Però ovviamente il discorso si può ripetere perché $K_{i_0} = K_{i_0 0} \cup \dots \cup K_{i_0(N-1)}$ per cui esiste un $i_1 \in \{0, \dots, N-1\}$ per cui $x \in K_{i_0 i_1}$. Per induzione si prosegue, costruendo un $\mathbf{i} \in N^\omega$ tale che $x \in \bigcap_{n \geq 0} K_{\mathbf{i}|n}$. In tale intersezione ci sta anche $\pi(\mathbf{i})$ per definizione. Ricordando però che quella intersezione è un singoletto, deve valere $x = \pi(\mathbf{i})$.

Verifichiamo (7). Siano $x = \pi(\mathbf{i})$ ed $\varepsilon > 0$. Allora per ogni n si ha $x \in K_{\mathbf{i}|n}$, ed inoltre esiste n_0 per cui $K_{\mathbf{i}|n_0} \subseteq B(x, \varepsilon)$ (sempre per il fatto che il diametro va a 0). Ne segue immediatamente che $[\mathbf{i} \upharpoonright n_0] \subseteq \pi^{-1}(B(x, \varepsilon))$ che è proprio la definizione di continuità. Osserviamo che la continuità di π è una ulteriore conferma del fatto che K è compatto, in quanto immagine di N^ω (compatto) tramite π (continua).

Passando a (4), fissiamo w e osserviamo che per definizione $s_w = S_w s_w$, quindi

$$K \supseteq K_w \supseteq K_{ww} \supseteq \dots \quad \text{e} \quad \pi(\bar{w}) \in \bigcap_{n \geq 0} K_{w^n},$$

ma tale intersezione contiene anche s_w in quanto punto fisso. Essendo essa un singoletto, ne segue che $s_w = \pi(\bar{w})$. Inoltre, per ogni n vale $\pi(\mathbf{i}) \in K_{\mathbf{i}|n} \ni s_{\mathbf{i}|n}$ per cui al limite $s_{\mathbf{i}|n} \xrightarrow{n \rightarrow \infty} \pi(\mathbf{i})$.

Ora vediamo (5) verificando i due contenimenti. Il verso (\supseteq) è ovvio perché K essendo compatto è anche chiuso e per il punto (4) tutti i punti fissi "finiti" stanno in K . L'altro segue sempre da (4) perché ogni $x = \pi(\mathbf{i})$ è limite dei vari $s_{\mathbf{i}|n}$, cioè ogni punto di K è di accumulazione per punti fissi "finiti".

Per verificare (6), dati \mathbf{i} e w si ha

$$\begin{aligned} K_{\mathbf{i}|0} &\supseteq K_{\mathbf{i}|1} \supseteq \dots \xrightarrow{\cap} \{\pi(\mathbf{i})\} \\ S_w K_{\mathbf{i}|0} &\supseteq S_w K_{\mathbf{i}|1} \supseteq \dots \xrightarrow{\cap} S_w \{\pi(\mathbf{i})\} \end{aligned}$$

e quest'ultima catena corrisponde esattamente a

$$K_{w\mathbf{i}|0} \supseteq K_{w\mathbf{i}|1} \supseteq \dots \xrightarrow{\cap} \{\pi(w\mathbf{i})\}.$$

Vediamo la seconda parte di (8). Sia $H \in \mathcal{K}$. Abbiamo allora

$$\begin{aligned} d(\pi(\mathbf{i}), H_{\mathbf{i}|n}) &\leq d(\pi(\mathbf{i}), K_{\mathbf{i}|n}) + d(K_{\mathbf{i}|n}, H_{\mathbf{i}|n}) \leq \\ &\leq \text{diam}(K_{\mathbf{i}|n}) + r^n d(K, H) \leq \\ &\leq r^n (\text{diam}K + d(K, H)) \xrightarrow[\text{indip. da } \mathbf{i}]{n \rightarrow \infty} 0. \end{aligned}$$

Infine per verificare (8') osserviamo che sicuramente $H \supseteq \Phi(H) \supseteq \Phi^2(H) \supseteq \dots$ e l'intersezione è punto fisso per Φ . Poiché il punto fisso di Φ è unico, esso deve essere K . Ora, preso \mathbf{i} , vale la catena di implicazioni

$$H \supseteq K \implies H_{\mathbf{i}|1} \supseteq K_{\mathbf{i}|1} \implies \dots$$

Si ha quindi

$$\pi(\mathbf{i}) = \bigcap_{n \geq 1} K_{\mathbf{i}|n} \subseteq \bigcap_{n \geq 1} H_{\mathbf{i}|n}$$

ma quest'ultimo è un singoletto, per cui $\pi(\mathbf{i}) = \bigcap_{n \geq 1} H_{\mathbf{i}|n}$ come volevamo. □

2.7 Misura di Hausdorff

Definizione 19. Fissato $s \geq 0$ e $\delta > 0$, la famiglia $\{E_i\}_{i < \omega}$ è un δ -ricoprimento di $A \subseteq \mathbb{R}^n$ se

- $A \subseteq \bigcup_{i < \omega} E_i$;
- $\text{diam}E_i \leq \delta, \forall i < \omega$.

Proposizione 13. La funzione

$$(H_\delta^s)^*(A) := \inf \left\{ \sum_{i < \omega} \alpha_s \left(\frac{\text{diam}E_i}{2} \right)^s : \{E_i\}_{i < \omega} \text{ è un } \delta\text{-ricoprimento di } A \right\}$$

è una misura esterna.

Dimostrazione. Che valga $(H_\delta^s)^*(\emptyset) = 0$ è ovvio. Verifichiamo ora la condizione (2b) della definizione. Osserviamo che se $A \subseteq \bigcup_{i < \omega} B_i$, un δ -ricoprimento di $\bigcup_{i < \omega} B_i$ è anche un δ -ricoprimento di A . Allora vale

$$(H_\delta^s)^*(A) \leq \sum_{i < \omega} (H_\delta^s)^*(B_i)$$

perché al termine sinistro “si può solo togliere roba” (nel caso in cui due o più dei B_i si sovrappongono). □

Osservazione 20. Se $0 < \delta' \leq \delta$, allora $(H_{\delta'}^s)^*(A) \geq (H_{\delta}^s)^*(A)$ per ogni insieme A . Infatti con meno ricoprimenti a disposizione l'estremo inferiore che definisce tale misura non può diminuire. Poiché al diminuire di δ abbiamo una successione crescente, possiamo prenderne il limite per $\delta \rightarrow 0$ ottenendo la prossima definizione.

Definizione 21. La misura

$$(\mathcal{H}^s)^*(A) := \lim_{\delta \rightarrow 0} (H_{\delta}^s)^*(A) = \inf_{\delta > 0} (H_{\delta}^s)^*(A)$$

si dice *misura esterna s -dimensionale di Hausdorff*. Di conseguenza la sua restrizione alla σ -algebra \mathcal{F} degli insiemi misurabili, $\mathcal{H}^s := (\mathcal{H}^s)^*|_{\mathcal{F}}$ è detta *misura s -dimensionale di Hausdorff*.

Lemma 14. Se $\mathcal{H}^s(A) < \infty$ e $t > s$, allora $\mathcal{H}^t(A) = 0$.

Dimostrazione. Sia $\{E_i\}_{i < \omega}$ un δ -ricoprimento di A . Allora per ogni i vale

$$(\text{diam} E_i)^t = (\text{diam} E_i)^{t-s} (\text{diam} E_i)^s \leq \delta^{t-s} (\text{diam} E_i)^s,$$

per cui passando all'estremo inferiore

$$\mathcal{H}_{\delta}^t(A) \leq \delta^{t-s} \mathcal{H}_{\delta}^s(A).$$

Ora, dato che per ipotesi $\mathcal{H}^s(A) < \infty$, facendo tendere $\delta \rightarrow 0$ il termine a destra va a zero e quindi si ottiene $\mathcal{H}^t(A) = 0$. \square

Osservazione 22. Questo lemma ha delle conseguenze interessanti:

1. per la contronominale all'enunciato, $\mathcal{H}^t(A) > 0$ e $s < t$ implicano $\mathcal{H}^s(A) = \infty$;
2. $\sup\{s : \mathcal{H}^s(A) = +\infty\} = \inf\{t : \mathcal{H}^t(A) = 0\} =: \dim_H(A)$ è detta *dimensione di Hausdorff di A* .

Definizione 23. Se A è boreliano, ha dimensione $\dim_H(A) = s$ e vale $\mathcal{H}^s(A) \neq 0, \infty$ allora diremo che A è un s -insieme.

2.8 Dimensione box

È possibile considerare una differente costruzione della misura. Supponiamo di considerare, invece di ricoprimenti con diametro $\leq \delta$, solo i ricoprimenti con diametro esattamente δ . Allora la misura di Hausdorff si può scrivere come

$$\mathcal{H}_{\delta}^s(A) = \inf\left\{\sum (\text{diam} E_i)^s : \dots\right\} = \delta^s N_{\delta}(A)$$

dove $N_{\delta}(A)$ è il minimo numero di insiemi di diametro δ che occorrono per ricoprire A . A differenza della costruzione fatta prima non si può scrivere $\lim_{\delta \rightarrow 0} \delta^s N_{\delta}(A)$ alla buona, perché non è affatto chiaro che $\delta^s N_{\delta}(A)$ cresca! Ad ogni modo si dà la seguente

Definizione 24. Si dice che A ha *dimensione box* $s \geq 0$ se il limite

$$\lim_{\delta \rightarrow 0} \delta^s N_{\delta}(A)$$

esiste diverso da $0, \infty$.

Si può esplicitare s in modo da ottenerlo tramite calcolo, se il limite scritto sopra esiste. Infatti passando ai logaritmi, deve valere

$$\log N_\delta(A) + s \log \delta \rightarrow \text{cost},$$

e dividendo per $-\log \delta$ abbiamo

$$\frac{\log N_\delta(A)}{-\log \delta} \rightarrow s =: \dim_B(A).$$

Per ovviare al problema di definizione che si pone quando il limite non esiste, a partire da quanto appena ricavato si definiscono anche altre due quantità.

Definizione 25. Definiamo *dimensione box inferiore* e *superiore* di A rispettivamente le due quantità

$$\begin{aligned} \underline{\dim}_B(A) &:= \liminf_{\delta \rightarrow 0} \frac{\log N_\delta(A)}{-\log \delta} \\ \overline{\dim}_B(A) &:= \limsup_{\delta \rightarrow 0} \frac{\log N_\delta(A)}{-\log \delta} \end{aligned}$$

Naturalmente quando esse coincidono sono anche uguali alla dimensione box dell'insieme. Nel libro di Falconer si dimostra che è possibile sostituire il numero $N_\delta(A)$ indifferentemente con:

- il numero minimo di insiemi di diametro $\leq \delta$ che occorre per ricoprire A ;
- il numero minimo di palle chiuse di raggio esattamente δ che occorre per ricoprire A ;
- il numero minimo di cubi chiusi di lato esattamente δ che occorre per ricoprire A ;
- il numero minimo di cubi di lato esattamente δ disposti a griglia che occorre per ricoprire A ;
- il numero *massimo* di palle aperte *disgiunte* di raggio δ con centro in A .

2.9 Proprietà di \dim_H e \dim_B

1. $A \subseteq B \implies \dim_H(A) \leq \dim_H(B)$.

Basta dimostrare che $\forall s, \mathcal{H}^s(A) = +\infty \implies \mathcal{H}^s(B) = +\infty$, e questo è ovvio per la definizione di misura (esterna).

2. Se $A \subseteq \mathbb{R}^n$ e $s > n$ allora $\mathcal{H}^s(A) = 0$ (e quindi $\dim_H(A) \leq n$).

Infatti senza perdita di generalità possiamo supporre $A \subseteq [0, 1]^n$ (se così non fosse, basta ricoprire A di cubetti e considerare quelli; la additività della misura fa il resto). In particolare A è copribile con una mesh di lato δ , quindi

$$\mathcal{H}^s(A) \leq \left(\frac{1}{\delta}\right)^n \cdot (\text{cost} \cdot \delta)^s = \text{cost}^s \cdot \delta^{s-n} \xrightarrow{\delta \rightarrow 0} 0.$$

3. $\dim_H \left(\bigcup_{i < \omega} A_i \right) = \sup_{i < \omega} \dim_H A_i$.

Infatti la disuguaglianza \geq è ovvia per il punto 1. Ora supponiamo per assurdo che valga $>$, per cui esiste un elemento separatore $\dim_H \left(\bigcup_{i < \omega} A_i \right) > s > \sup_{i < \omega} \dim_H A_i$. Per la definizione di dimensione di Hausdorff si ha quindi $\mathcal{H}^s(A_i) = 0$ per ogni i , da cui per la subadditività della misura segue

$$\mathcal{H}^s \left(\bigcup_{i < \omega} A_i \right) \leq \sum_{i < \omega} \underbrace{\mathcal{H}^s(A_i)}_{=0} = 0$$

ma allora dovrebbe essere $\dim_H \left(\bigcup_{i < \omega} A_i \right) = 0$, assurdo.

4. In particolare, se A è numerabile la sua dimensione di Hausdorff vale 0.

Basta ricoprirlo di singoletti, che hanno ovviamente dimensione di Hausdorff nulla.

Definizione 26. Una funzione $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ si dice *hölderiana* di costanti $\alpha, r > 0$ se

$$\|Fx - Fy\| \leq r\|x - y\|^\alpha \quad \forall x, y \in \mathbb{R}^n.$$

Definizione 27. Una funzione $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ si dice *bilipschitziana* di costante $r > 1$ se

$$r^{-1}\|x - y\| \leq \|Fx - Fy\| \leq r\|x - y\| \quad \forall x, y \in \mathbb{R}^n.$$

Proposizione 15. a) Sia F α -Hölderiana con $\alpha, r > 0$. Allora

1. $\forall s, \mathcal{H}^{s/\alpha}(FA) \leq r^{s/\alpha} \mathcal{H}^s(A);$
2. $\dim_H(FA) \leq \alpha^{-1} \dim_H(A)$

b) Se F è bilipschitziana allora $\dim_H(FA) = \dim_H(A)$.

c) Ogni F lipschitziana non fa crescere $\underline{\dim}_B$ e $\overline{\dim}_B$. Se è bilipschitziana allora le lascia invariate.

Dimostrazione. Dimostriamo per prima (a.1). Sia $\{E_i\}$ un δ -ricoprimento di A . Si ha quindi $A = \bigcup_i (A \cap E_i)$, per cui dato che F distribuisce sulle unioni si ottiene $FA = \bigcup_i F[A \cap E_i]$. Ora osserviamo che per ogni i vale, per α -Hölderianità,

$$\text{diam}(F[A \cap E_i]) \leq r[\text{diam}(A \cap E_i)]^\alpha \stackrel{(*)}{\leq} r(\text{diam } E_i)^\alpha \leq r\delta^\alpha.$$

Da ciò si ottengono due cose:

- $\{F[A \cap E_i]\}$ è un $r\delta^\alpha$ -ricoprimento di FA ;
- $[\text{diam}(F[A \cap E_i])]^{s/\alpha} \stackrel{(*)}{\leq} r^{s/\alpha} [\text{diam } E_i]^s$.

Quindi abbiamo ricoperto A e FA con lo stesso numero di insiemi, opportunamente riscalati. Pertanto otteniamo che

$$\mathcal{H}_{r\delta^\alpha}^{s/\alpha}(FA) \leq r^{s/\alpha} \mathcal{H}_\delta^s(A).$$

Poiché il ricoprimento $\{E_i\}$ di partenza è arbitrario, per $\delta \rightarrow 0$ si ha la tesi.

Per dimostrare (a.2) basta vedere che $s > \alpha^{-1} \dim_H(A) \implies \mathcal{H}^s(FA) = 0$, e questo viene dal punto precedente. Infatti da

$$\forall \tilde{s}, \quad \mathcal{H}^{\tilde{s}/\alpha}(FA) \leq r^{\tilde{s}/\alpha} \mathcal{H}^{\tilde{s}}(A)$$

ponendo $\tilde{s} = \alpha s$ otteniamo

$$\mathcal{H}^s(FA) \leq r^s \mathcal{H}^{s\alpha}(A).$$

Per ipotesi abbiamo $\alpha s < \dim_H(A)$, per cui $\mathcal{H}^{s\alpha}(A) = 0$ e di conseguenza anche il membro sinistro della disuguaglianza deve essere nullo come volevamo dimostrare.

Per dimostrare (b) si applica (a.2), forti del fatto che F è iniettiva (viene dalla bilipschitzianità) e quindi esiste F^{-1} :

$$\dim_H A \leq \dim_H F^{-1}FA \leq \dim_H FA \leq \dim_H A$$

per cui essendo il primo e l'ultimo termine uguali, valgono tutte come uguaglianze.

Infine dimostriamo (c) riprendendo il ragionamento di (a). In questo caso $\alpha = 1$, per cui se abbiamo un δ -ricoprimento di A dato da $\{E_i\}$, si ha che $\{F[A \cap E_i]\}$ è un $r\delta$ -ricoprimento di FA , per cui per ogni $\delta > 0$ si ha $N_{r\delta}(FA) \leq N_\delta(A)$. Passando i logaritmi e dividendo per $-\log \delta = -\log(r\delta) + \log r$, si ottiene (per $\delta < 1$, ma dovendo fare il limite questa condizione è soddisfatta definitivamente)

$$\frac{\log(N_{r\delta}(FA))}{-\log(r\delta) + \log r} \leq \frac{\log(N_\delta(A))}{-\log \delta}$$

e questa disuguaglianza vale passando al \liminf e al \limsup per $\delta \rightarrow 0$, per cui dato che il $\log r$ viene mangiato dagli altri termini si ottengono

$$\underline{\dim}_B(FA) \leq \underline{\dim}_B(A), \quad \overline{\dim}_B(FA) \leq \overline{\dim}_B(A).$$

□

Ulteriori conseguenze:

- Ogni affinità non singolare di \mathbb{R}^n , essendo bilipschitziana, conserva \dim_H e \dim_B .
- Se \mathcal{C} è una varietà liscia di dimensione m immersa in \mathbb{R}^n tramite F , allora essendo F bilipschitziana vale $\dim_H \mathcal{C} = m$.
- Se $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è un'affinità (eventualmente singolare, come può essere una proiezione) allora essa è lipschitziana e quindi $\dim_H(FA) \leq \min\{\dim_H(A), m\}$.

C'è un aspetto *apparentemente* gradevole di \dim_B : per ogni $A \subseteq \mathbb{R}^n$ limitato, si hanno

$$\underline{\dim}_B A = \underline{\dim}_B A^f, \quad \overline{\dim}_B A = \overline{\dim}_B A^f,$$

dove al solito f indica la chiusura topologica di un insieme. Questo significa che se consideriamo l'insieme $[0, 1]^2 \cap \mathbb{Q}^2$, questo ha dimensione di Hausdorff nulla in quanto è numerabile, ma d'altro canto ha la stessa dimensione box della sua chiusura, cioè 2!

Dimostrazione (dimensione box e chiusura). Ricordiamo che per definizione

$$\underline{\dim}_B(A) = \liminf_{\delta \rightarrow 0} \frac{\log N_\delta(A)}{-\log \delta},$$

dove in questo caso nel calcolo di $N_\delta(A)$ scegliamo di utilizzare palle chiuse di raggio esattamente δ . Poiché $[0, 1]^2 \times \mathbb{Q}^2$ è denso in $[0, 1]^2$, coprire il primo è esattamente come coprire il secondo, per qualunque scelta di $\delta > 0$. Pertanto facendo il limite si ottiene lo stesso numero considerando $N_\delta(A)$ o $N_\delta(A^f)$. Analogamente si dimostra l'asserzione riguardo alla dimensione box superiore. \square

2.10 Tecniche per calcolare $\dim_H A$

Solitamente, a meno che non si sia in un caso fortunato, è difficile calcolare la dimensione di Hausdorff di un oggetto. Quello che si cerca di fare è dare delle stime superiori e inferiori a tale quantità. Presentiamo ora qualche risultato che ci aiuta in questa operazione.

Lemma 16. *Vale $\dim_H A \leq \underline{\dim}_B A \leq \overline{\dim}_B A$.*

Dimostrazione. La seconda disuguaglianza è ovvia. Per la prima, osserviamo preliminarmente che per $\dim_H A = 0$ è ovvia. Possiamo quindi supporre che esista $s > 0$ per cui $\mathcal{H}^s(A) = +\infty$. Allora per $\delta > 0$ abbastanza piccolo (in particolare minore di 1), abbiamo

$$1 < \mathcal{H}_\delta^s(A) \leq N_\delta(A)\delta^s$$

di cui è importante la disuguaglianza a destra, che segue dal fatto che gli insiemi usati per calcolare $\mathcal{H}_\delta^s(A)$ sono stati tutti “gonfiati” fino a raggiungere il loro massimo diametro δ . Passando ai logaritmi otteniamo quindi

$$0 < \log N_\delta(A) + s \log \delta,$$

da cui

$$s < \frac{\log N_\delta(A)}{-\log \delta}$$

per cui passando al $\liminf_{\delta \rightarrow 0}$, per la permanenza del segno vale $s \leq \underline{\dim}_B A$. Infine, essendo s arbitrario possiamo passare al \sup_s , ottenendo

$$\dim_H A \leq \underline{\dim}_B A.$$

\square

Definizione 28. Diremo che un insieme A è *totalmente sconnesso* se per ogni $x, y \in A$ esiste un clopen di A che li separa.

Lemma 17. *Se $\dim_H A < 1$ allora A è totalmente sconnesso.*

Dimostrazione. Prima di dimostrare il lemma vero e proprio, facciamo un'osservazione. Fissato $x_0 \in \mathbb{R}^n$, la funzione $F(x) := \|x - x_0\|$ è lipschitziana di costante 1 grazie alla disuguaglianza triangolare:

$$\left| \|x - x_0\| - \|y - y_0\| \right| \leq \|x - y\| \implies |Fx - Fy| \leq \|x - y\|.$$

Ora consideriamo due generici x_0 e x_1 punti distinti di A e la F appena definita. Abbiamo che $x_0 \mapsto 0$ e $x_1 \mapsto d > 0$. Da quanto visto in precedenza sul legame tra dimensione di Hausdorff e funzioni lipschitziane, la dimensione dell'immagine non può crescere, per cui vale ancora $\dim_H F[A] < 1$. Quindi $F[A]$ non può contenere alcun intervallo, perché in quel caso la dimensione dovrebbe essere almeno 1 per la proprietà 1 di \dim_H . In particolare non può contenere l'intervallo $[0, d]$, per cui esiste un certo $0 < r < d$ che non sta in $F[A]$. Se consideriamo quindi la palla $B(x_0, r)$, essa non ha punti di A sul bordo. Per questo motivo se la intersechiamo con A otteniamo un clopen di A che contiene x_0 e il cui complementare contiene x_1 , per cui abbiamo effettivamente separato x_0 e x_1 . A è quindi totalmente sconnesso. \square

Teorema 18. *Sia $A \subseteq \mathbb{R}^n$ limitato. Supponiamo di avere, per ogni $n \geq 0$, un ricoprimento di A con $N(n)$ insiemi di diametro al più $\delta(n)$. Assumiamo inoltre che $\delta(n) \xrightarrow{n \rightarrow +\infty} 0$ non superesponenzialmente, ossia $\exists c \in (0, 1)$ t.c. $\delta(n+1) \geq c \cdot \delta(n) \forall n \geq 0$. Allora:*

1.

$$\underline{\dim}_B A \leq \liminf_{n \rightarrow \infty} \frac{\log N(n)}{-\log \delta(n)};$$

2.

$$\overline{\dim}_B A \leq \limsup_{n \rightarrow \infty} \frac{\log N(n)}{-\log \delta(n)};$$

3. *se per un fissato $s \in \mathbb{R}_{\geq 0}$ l'insieme $\{N(n) \cdot \delta(n)^s\}_n$ è limitato, allora $\mathcal{H}^s(A) < \infty$ (e dunque $\dim_H A \leq s$).*

Prima di dare la dimostrazione applichiamo il teorema con qualche esempio familiare.

Esempio 3. • **Insieme di Cantor standard.** In questo caso possiamo, pensando al primo passo della costruzione, dedurre che vanno bene $N(n) = 2^n$ e $\delta(n) = (1/3)^n$. Otteniamo quindi

$$\frac{\log 2^n}{-\log (1/3)^n} = \frac{\log 2}{\log 3}.$$

• **Triangolo di Sierpinski.** Se consideriamo la costruzione classica fatta nel quadrato unitario, con $N(n) = 3^n$, $\delta(n) = \sqrt{2}(1/2)^n$ otteniamo facilmente la stima $\log 3 / \log 2$.

• **Curva di Koch.** In questo caso delle buone scelte sono $N(n) = 4^n$ e $\delta(n) = (1/3)^n$, per cui la stima che si ottiene è $\log 4 / \log 3$.

Dimostrazione. Sia $\delta > 0$ arbitrario, scelto in modo che $\delta(n+1) \leq \delta \leq \delta(n)$. Allora per la monotonia del logaritmo, e dall'ipotesi sulla crescita di $\delta(n)$, vale

$$\frac{\log N_\delta(A)}{-\log \delta} \leq \frac{\log N_{\delta(n+1)}(A)}{-\log \delta(n)} \leq \frac{\log N_{\delta(n+1)}(A)}{-\log \delta(n+1) + \log c}.$$

Di conseguenza vale, dato che al limite $\log c$ non influisce essendo costante,

$$\limsup_{\delta \rightarrow 0} \frac{\log N_\delta(A)}{-\log \delta} \leq \limsup_{n \rightarrow \infty} \frac{\log N_{\delta(n)}(A)}{-\log \delta(n)}.$$

Poiché l'insieme $\{\delta(n)\}_n$ su cui facciamo il limite a destra è contenuto in quello dei δ di sinistra, vale anche la disuguaglianza opposta e di conseguenza l'uguaglianza. Pertanto nel calcolo di $\overline{\dim}_B$ si possono utilizzare le successioni $\delta(n)$. Si dimostra analogamente questo fatto anche per \liminf e $\underline{\dim}_B$.

Applicando quanto appena dimostrato vale quindi

$$\overline{\dim}_B A = \limsup_{n \rightarrow \infty} \frac{\log N_{\delta(n)}(A)}{-\log \delta(n)} \leq \limsup_{n \rightarrow \infty} \frac{\log N(n)}{-\log \delta(n)}$$

dove la disuguaglianza vale perché a sinistra consideriamo la cardinalità del $\delta(n)$ -ricoprimento minimo di A , mentre a destra è la cardinalità di un $\delta(n)$ -ricoprimento arbitrario. Abbiamo quindi dimostrato il punto 2. Analogamente si dimostra il punto 1.

Vediamo l'affermazione 3. Supponiamo quindi che per ogni n valga $N(n)\delta(n)^s \leq M < \infty$ per un certo M . Allora per definizione vale anche $\mathcal{H}_{\delta(n)}^s(A) \leq M$. Per $n \rightarrow \infty$ si ottiene che $\mathcal{H}^s(A) \leq M$, per cui abbiamo concluso. \square

Proposizione 19 (principio della poca massa). *Sia μ una misura finita su \mathbb{R}^n con $\text{supp}(\mu) \subseteq A$. Sia $s > 0$ per cui $\exists c, \delta > 0$ t.c. $\forall E$ con $\text{diam } E \leq \delta$ si ha $\mu(E) \leq c(\text{diam } E)^s$. Allora $\mathcal{H}^s(A) \geq c^{-1}\mu(A)$ (in particolare $\epsilon > 0$, per cui $\dim_H A \geq s$).*

Dimostrazione. Sia $\{E_i\}_{i < \omega}$ un δ -ricoprimento di A . Allora per definizione di misura e per ipotesi,

$$0 < \mu(A) \leq \sum_{i < \omega} \mu(E_i) \leq c \cdot \sum_{i < \omega} (\text{diam } E_i)^s.$$

Dunque prendendo l'estremo inferiore sui vari δ -ricoprimenti di A , si ottiene

$$\mu(A) \leq c\mathcal{H}_\delta^s(A)$$

per cui per $\delta \rightarrow 0$ si ha la tesi. \square

Esempio 4. Consideriamo l'insieme di Cantor standard, denotandolo con K . Il modo intuitivo per metterci una misura è quello di considerare lo spazio $(2^\omega, P_{(1/2, 1/2)})$ e utilizzare la applicazione di proiezione $\pi : 2^\omega \rightarrow K$, definendo la misura pushforward su K : $\mu = \pi_* P$. Consideriamo dunque $s = \log 2 / \log 3$, $\delta = 1$, $c = 2$, e come insieme tipico di un ricoprimento di K consideriamo senza perdita di generalità un intervallo E . Quale che sia la sua lunghezza, esiste certamente un $n \geq 0$ per cui $(1/3)^{n+1} \leq \text{diam } E < (1/3)^n$. Allora segue che

$$\mu(E) \leq \left(\frac{1}{2}\right)^n = \left(\frac{1}{3}\right)^{sn} = 3^s \left(\frac{1}{3}\right)^{s(n+1)} \leq 2(\text{diam } E)^s.$$

Quindi $\dim_H A \geq s = \log 2 / \log 3$, per cui unito al risultato ottenuto prima questo ci dice che tale numero è esattamente la dimensione di Hausdorff dell'insieme di Cantor standard.

Esercizio 1. Considerato l'IFS definito dalle due trasformazioni

$$Q_0 : \begin{pmatrix} r & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q_1 : \begin{pmatrix} 1-r & 0 & r \\ 0 & 1-p & p \\ 0 & 0 & 1 \end{pmatrix}$$

trovare l'attrattore (numericamente) e stimarne la dimensione di Hausdorff. Dimostrare che è un omeomorfismo, che è derivabile λ -q.o. (suggerimento: teorema di Lebesgue), e che dove la derivata esiste questa è nulla (giustificando la classificazione dell'attrattore come *scala del diavolo*).

Soluzione. Consideriamo lo spazio $(2^\omega, P_{(r,1-r)})$ e la mappa $\pi : 2^\omega \rightarrow K$ dove K è l'attrattore "orizzontale", e la misura push-forward $\mu = \pi_*P$. L'attrattore dell'IFS grande è la funzione di ripartizione M . Essa è continua in quanto per ogni \mathbf{i} vale $\mu(\{\pi(\mathbf{i})\}) = 0$ perché il termine di sinistra non è altro che $P(\pi^{-1}\{x\})$ e l'insieme all'interno è composto al più di due punti. Inoltre è iniettiva perché dati $x < x'$ si può sempre trovare una parola w tale che $S_w[0, 1] \subset [x, x']$ e la massa contenuta in tale intervallo è $(1/2)^{\#w}$, per cui $M(x) < M(x')$.

Da queste due proprietà segue per il Teorema di Lebesgue che la funzione M è derivabile λ -q.o. Ciò che rimane da dimostrare è contenuto nel seguente teorema. \square

Teorema 20. Se $r \neq 1/2$ allora M è differenziabile λ -q.o. e la derivata vale 0.

Osservazione 29. Questo qualifica la funzione M come scala del diavolo *scivolosa*, in contrapposizione con scale del diavolo non scivolose in cui i gradini sono effettivamente orizzontali.

Osservazione 30. Se consideriamo il processo stocastico i.i.d. $X = (X_n, n \geq 0)$ con $X_n : (2^\omega, P_{(r,1-r)}) \rightarrow \{0, 1\}$ che indica il valore della n -esima lettera di una parola infinita \mathbf{i} e applichiamo la legge forte dei grandi numeri, otteniamo che

$$\frac{\#\{1 \text{ che compaiono fino al livello } n\}}{n} \xrightarrow{\text{q.o.}} 1 - r.$$

Dimostrazione. Abbiamo osservato che $\pi_*P_{1/2} = \lambda$, dunque poichè un tipico elemento $x = \pi(\mathbf{i})$ è incapsulato in un rettangolo centrato in $(x, M(x))$ con base $b(n) = S_{\mathbf{i}|n}([0, 1])$ e altezza $a(n) = T_{\mathbf{i}|n}([0, 1])$, per l'osservazione precedente abbiamo per la derivata

$$M'(x) = \lim_{n \rightarrow \infty} \frac{a(n)}{b(n)} = \lim_{n \rightarrow \infty} \frac{(1/2)^n}{r^{nr}(1-r)^{n(1-r)}} = \lim_{n \rightarrow \infty} \left(\frac{1}{2r^r(1-r)^{1-r}} \right)^n.$$

A questo punto il comportamento del limite dipende dal confronto tra la quantità tra parentesi e 1. Introduciamo il risultato che ci permette di concludere la dimostrazione. \square

Teorema 21. Sia $h : \Delta = \{(x_0, \dots, x_{N-1}) \text{ vettori di prob.}\} \rightarrow \mathbb{R}_{\geq 0}$ definita da

$$h(\mathbf{x}) = x_0(-\log x_0) + \dots + x_{N-1}(-\log x_{N-1}).$$

Essa è anche nota come *entropia metrica dello shift pieno su N simboli rispetto alla metrica prodotto $P_{\mathbf{x}}$* . Allora h ha uno ed un solo massimo in $(1/N, \dots, 1/N)$ e tale massimo è $\log N$.

Dimostrazione. L'abbiamo vista numericamente per vettori 3-dimensionali. Si calcola il massimo usando tecniche di Analisi II (jacobiano, hessiano). Attenzione, bisogna sfruttare il fatto che si tratta di un vettore di probabilità: $x_{N-1} = -(x_0 + \dots + x_{N-2})$, per cui in realtà h dipende solo dai primi $N - 1$ valori di \mathbf{x} . \square

Fine dimostrazione del teorema. Per $N = 2$ abbiamo il caso che ci serve, infatti si ottiene

$$r(-\log r) + (1-r)(-\log(1-r)) \leq \log 2$$

cioè

$$r \log r + (1-r) \log(1-r) \geq \log \frac{1}{2}$$

con minimo in $r = 1/2$, per cui applicando exp si trova

$$r^r(1-r)^{1-r} \geq \frac{1}{2}.$$

Per cui in conclusione se $r \neq 1/2$, la quantità nel limite è più piccola di 1 quindi $M'(x) = 0$. \square

2.11 Misure stazionarie

Sia $H \subseteq \mathbb{R}^n$ compatto. Indichiamo $\mathcal{M}(H)$ l'insieme delle misure finite boreliane su H e $\mathcal{P}(H)$ l'insieme delle probabilità boreliane su H . Entrambi questi insiemi sono dotati della topologia debole data considerando l'inclusione $\mathcal{M}(H) \hookrightarrow \mathbb{R}^{\mathcal{C}(H)}$ che manda $\mu \mapsto \int_H f d\mu$. Infatti

$$\mu_n \rightarrow \mu \iff \forall f \in \mathcal{C}(H), \int_H f d\mu_n \rightarrow \int_H f d\mu.$$

Sia ora $\mathcal{S} = \{S_0, \dots, S_{N-1}\}$. Abbiamo visto l'operatore di Hutchinson $\Phi : \mathcal{K}(H) \rightarrow \mathcal{K}(H)$. Consideriamo ora anche, per $\mathbf{p} \in \Delta$, $\mathbf{p} = (p_0, \dots, p_{N-1})$, l'operatore su $\mathcal{M}(H)$ (che chiameremo sempre operatore di Hutchinson) dato da

$$\Phi_{\mathbf{p}}(\nu) = \sum_{i=0}^{N-1} p_i(S_i)_*(\nu).$$

Esercizio 2. Sia S_* che $\Phi_{\mathbf{p}}$ distribuiscono sulle somme positive e sui prodotti per scalari positivi.

Quindi applicando l'operatore ad una probabilità, per come abbiamo scelto i pesi p_i il risultato è sempre una probabilità. I punti fissi per $\Phi_{\mathbf{p}}$ vengono detti *probabilità stazionarie*.

Definizione 31. La *distanza di 1-Wasserstein* su $\mathcal{P}(H)$ è

$$W(\mu, \nu) = \sup \left\{ \int_H f d\mu - \int_H f d\nu : f : H \rightarrow \mathbb{R} \text{ 1-Lipschitz} \right\}.$$

Teorema 22. La distanza W metrizza la topologia debole.

Una definizione alternativa della distanza di 1-Wasserstein è dovuta a Kantorovich, ed è

$$W(\mu, \nu) = \inf \left\{ \int_{H^2} \|x - y\| d\rho(x, y) : \rho \text{ è un accoppiamento tra } \mu \text{ e } \nu \right\},$$

dove si intende dire che μ e ν sono i marginali di ρ . Questa definizione è interessante anche perché si lega ad un problema concreto, quello del trasporto ottimo in cui il costo per il trasporto è pari alla distanza tra i due punti da congiungere.

Vediamo alcune proprietà della distanza di 1-Wasserstein. Per prima cosa controlliamo che sia una vera distanza. L'unica condizione non banale da verificare è la disuguaglianza triangolare. Fissiamo quindi una f che sia 1-Lipschitz e osserviamo che per $\mu, \nu, \xi \in \mathcal{P}(H)$ vale

$$|\mu(f) - \nu(f)| \leq |\mu(f) - \xi(f)| + |\xi(f) - \nu(f)| \leq W(\mu, \xi) + W(\xi, \nu)$$

per cui passando all'estremo superiore sulle varie f si ottiene proprio quanto cercato.

Un'altra proprietà è il fatto che poiché H è compatto, vale $H \subseteq B(a, R)$ per opportuni $a \in \mathbb{R}^n$ e $R > 0$, e quindi la distanza W è limitata. Infatti prese $\mu, \nu \in \mathcal{P}(H)$ abbiamo

$$\mu(f) - \nu(f) = \mu(f - f(a) + f(a)) - \nu(f - f(a) + f(a)) = \mu(f - f(a)) + f(a) - \nu(f - f(a)) - f(a) \leq 2R$$

dove l'ultima disuguaglianza si ottiene dato che f è 1-Lipschitz e quindi $|f(x) - f(a)| \leq |x - a| \leq R$ per ogni x , e possiamo omettere il valore assoluto senza perdita di generalità.

Infine vediamo cosa succede calcolando la distanza W tra due misure degeneri, diciamo in a e in b . Abbiamo, sempre grazie alla 1-lipschitzianità, che

$$W(\delta_a, \delta_b) = \sup\{|f(a) - f(b)| : f \text{ è 1-Lips.}\} \leq \|a - b\|.$$

Teorema 23. Dato un IFS $S = \{S_0, \dots, S_{N-1}\}$ con $S_i : H \rightarrow H$ per ogni i di costanti $0 < r_1, \dots, r_{N-1} < 1$ e fissato $\mathbf{p} = (p_0, \dots, p_{N-1})$ vettore di probabilità abbiamo l'operatore di Hutchinson (per probabilità) $\Phi : \mathcal{P}(H) \rightarrow \mathcal{P}(H)$ definito da

$$\Phi(\mu) := \sum_{i=0}^{N-1} p_i (S_i)_*(\mu).$$

Allora Φ è una contrazione stretta rispetto alla metrica W . Quindi esiste unico il punto fisso μ e per ogni $\nu \in \mathcal{P}(H)$, vale $\Phi^n(\nu) \xrightarrow[n \rightarrow \infty]{\text{in } W} \mu$. In altre parole, se f è continua su H abbiamo

$$\int_H f d\mu \xleftarrow{n \rightarrow \infty} \int_H f d(\Phi^n \nu) = \int_H f d \left(\sum_{|w|=n} p_w \cdot (S_w)_*(\nu) \right) = \sum_{|w|=n} p_w \cdot \nu(f \circ S_w).$$

Tale unico punto fisso è proprio $\pi_*(P_{\mathbf{p}})$, e inoltre se K è l'attrattore dell'IFS vale $\text{supp}(\mu) = K$.

Dimostrazione. Sia $r = \max\{r_1, \dots, r_{N-1}\} < 1$. Fissiamo una f 1-lipschitziana e consideriamo $\mu, \nu \in \mathcal{P}(H)$. Allora

$$\begin{aligned} \left| \int f d(\Phi\mu) - \int f d(\Phi\nu) \right| &= \left| \sum p_i \mu(f \circ S_i) - \sum p_i \nu(f \circ S_i) \right| = \\ &= \left| \sum p_i r [\mu(r^{-1}f \circ S_i) - \nu(r^{-1}f \circ S_i)] \right| \stackrel{(*)}{\leq} \\ &\leq \underbrace{\sum p_i}_=1 r W(\mu, \nu) = r W(\mu, \nu) \end{aligned}$$

dove il passaggio (*) segue dal fatto che $r^{-1}f \circ S_i$ è 1-lipschitziana. Infatti presi x e y qualsiasi vale, dato che f è 1-lipschitziana,

$$|r^{-1}f(S_i(x)) - r^{-1}f(S_i(y))| \leq r^{-1}\|S_i(x) - S_i(y)\| = r^{-1}r_i\|x - y\| \leq \|x - y\|.$$

Passando quindi all'estremo superiore sulle f 1-lipschitziane nella disuguaglianza trovata, otteniamo

$$W(\Phi\mu, \Phi\nu) \leq r W(\mu, \nu), \quad 0 < r < 1,$$

che dimostra che Φ è una contrazione stretta di costante r rispetto a W .

Quindi sappiamo per il teorema sulle contrazioni strette che esiste un unico punto fisso μ per Φ . Per verificare che $\mu = \pi_*(P_{\mathbf{p}})$ basta quindi verificare che $\pi_*(P)$ (d'ora in poi ometteremo il pedice) è punto fisso per Φ . Osserviamo che, se $\sigma_i : N^\omega \rightarrow [i] \subseteq N^\omega$ è l'operatore di shift, ossia $\sigma_i(\mathbf{j}) = i\mathbf{j}$, abbiamo che

$$P \stackrel{\pm}{=} \sum_i p_i (\sigma_i)_* P.$$

Infatti per vederlo possiamo senza perdita di generalità considerare un blocco iniziale $A = [w]$ (generano tutta la σ -algebra!) e per esso vale

$$\sum_i p_i P(\sigma_i^{-1}[A]) = \sum_i p_i P(A|[i]) = P(A)$$

dove la seconda uguaglianza è contenuta nel teorema di Bayes e la prima segue dal fatto che ora dimostriamo che

$$P(\sigma_i^{-1}[w]) = P([w]|[i]).$$

Si hanno due casi:

1. la prima lettera in w non è i . Allora $\sigma_i^{-1} = \emptyset$, per cui la misura a sinistra è 0. D'altro canto anche $[w] \cap [i] = \emptyset$, per cui per definizione di probabilità condizionata anche la misura a destra è 0;
2. $w = iu$ per un'altra parola u . Abbiamo allora

$$\begin{aligned} P(\sigma_i^{-1}[w]) &= P([u]) \\ P([w]|[i]) &= \frac{P([iu] \cap [i])}{P([i])} = \frac{P([i])P([u])}{P([i])} = P([u]) \end{aligned}$$

dove abbiamo usato il fatto che P è una misura prodotto, per cui $P([iu]) = P([i])P([u])$.

Ora abbiamo che il diagramma

$$\begin{array}{ccc} N^\omega & \xrightarrow{\sigma_i} & N^\omega \\ \downarrow \pi & & \downarrow \pi \\ K & \xrightarrow{S_i} & K \end{array}$$

commuta perché dal teorema fondamentale degli IFS, preso $\mathbf{j} \in N^\omega$ si ha che $S_i(\pi(\mathbf{j})) = \pi(i\mathbf{j})$. Allora

$$\begin{aligned} \Phi(\pi_* P) &= \sum_i p_i (S_i)_* (\pi_* P) = \sum_i p_i (S_i \circ \pi)_* P = \sum_i p_i (\pi \circ \sigma_i)_* P = \\ &= \sum_i p_i \pi_* (\sigma_{i*} P) = \pi_* \left(\sum_i p_i \sigma_{i*} P \right) \stackrel{+}{=} \pi_* P. \end{aligned}$$

Resta da vedere che $\text{supp}(\mu) = K$. L'inclusione (\subseteq) è ovvia perché $\text{supp}(\mu) = \text{supp}(\pi_* P) \subseteq \pi(N^\omega) = K$, quindi verifichiamo il viceversa. Sia $x \in K$ e fissiamo $\varepsilon > 0$ arbitrario. Abbiamo dimostrato, quando abbiamo visto che π è continua, che $\pi^{-1}(B(x, \varepsilon))$ contiene un cilindro, diciamo $[w]$. Questo implica che $x \in \text{supp}(\mu)$. \square

2.12 Dimensione di similitudine

Consideriamo l'insieme di mappe S con le loro costanti r_i definito prima. D'ora in poi sia D la soluzione (unica!) di $\sum r_i^D = 1$ e poniamo $p_i = r_i^D$. Così facendo possiamo definire un vettore di probabilità $\mathbf{p} = (p_0, \dots, p_{N-1})$.

Osservazione 32. Per ogni $n \geq 0$ si ha che $\sum_{|w|=n} r_w^D = \sum_{|w|=n} p_w = 1$ in quanto $p_w = P([w])$.

Proposizione 24. 1. $\mathcal{H}^D(K) < \infty$, quindi $\dim_H K = s \leq D$;

2. Se le S_i sono **similitudini** e inoltre K è un s -insieme allora

$$s = D \iff \forall i \neq j, \mathcal{H}^s(K_i \cap K_j) = 0.$$

Dimostrazione. (1) Osserviamo che per ogni $n \geq 0$ l'unione dei vari K_w con $|w| = n$ copre K , e inoltre sempre per ogni n vale

$$\sum_{|w|=n} (\text{diam } K_w)^D \leq \sum_{|w|=n} (r_w \text{diam } K)^D = (\text{diam } K)^D \sum_{|w|=n} r_w^D = (\text{diam } K)^D.$$

Per $n \rightarrow \infty$ si ha $\text{diam } K_w \rightarrow 0$, quindi aspettando abbastanza a lungo otteniamo un ricoprimento di K fine a piacere, per cui $\mathcal{H}^D(K) \leq (\text{diam } K)^D < \infty$.

(2) (\Leftarrow) Si ha che, dato che il ricoprimento di K con i K_i è disgiunto (almeno rispetto a quanto rilevabile da \mathcal{H}^s), vale

$$\mathcal{H}^s(K) = \sum_i \mathcal{H}^s(K_i) = \sum_i r_i^s \mathcal{H}^s(K) = \mathcal{H}^s(K) \sum_i r_i^s,$$

quindi dato che K è un s -insieme, ossia $0 < \mathcal{H}^s(K) < \infty$, possiamo dividere per tale misura ottenendo $1 = \sum_i r_i^s$ ma dato che tale equazione ha soluzione unica che sappiamo essere D , vale $s = D$. (\Rightarrow) Si ha

$$\mathcal{H}^D(K) \leq \sum_i \mathcal{H}^D(K_i) = \sum_i r_i^D \mathcal{H}^D(K) = \mathcal{H}^D(K)$$

per cui in realtà vale l'uguaglianza, da cui segue che i K_i non possono avere intersezione di misura positiva. \square

Ricordiamo la

Definizione 33. Diciamo che il nostro IFS S soddisfa la *open set condition* (OSC) se esiste un aperto $O \neq \emptyset$ tale che:

1. $\Phi(O) \subseteq O$;
2. se $i \neq j$, allora $S_i(O) \cap S_j(O) = \emptyset$.

Osservazione 34. Se le due condizioni valgono per un qualche insieme *chiuso* H , allora valgono anche per un suo ingrossamento $H(\varepsilon)$ per un $\varepsilon > 0$ abbastanza piccolo. Infatti per un $\varepsilon > 0$ generico si ha

$$\begin{aligned} S_i[H(\varepsilon)] &= \bigcup_{x \in H} S_i[B(x, \varepsilon)] = \bigcup_{x \in H} B(S_i x, r_i \varepsilon) = \\ &= (S_i H)(r_i \varepsilon) \subseteq \left(\bigcup_i S_i H \right) (r_{\max} \varepsilon) \subseteq H(\varepsilon) \end{aligned}$$

dove l'ultima inclusione si ha perché $\cup_i S_i H \subseteq H$ e $r_{\max} < 1$. Inoltre se $i \neq j$, sia $d := \min_{i \neq j} d(H_i, H_j) > 0$ e prendiamo ε t.c. $r_{\max} \cdot \varepsilon < d/2$. Allora si ha

$$S_i[H(\varepsilon)] \cap S_j[H(\varepsilon)] = H_i(r_i \varepsilon) \cap H_j(r_j \varepsilon) = \emptyset.$$

Osservazione 35. Le S_i commutano con gli operatori topologici (interno, chiusura, complementare, bordo...).

Lemma 25. Supponiamo che S soddisfi la OSC rispetto a O . Allora

- (i) $\forall \mathbf{i} \in N^\omega$, $O \supseteq O_{\mathbf{i}|1} \supseteq O_{\mathbf{i}|2} \supseteq \dots$;
- (ii) $\forall w \in N^*$, $K_w \subseteq O_w^f$, compreso il caso particolare di $w = \emptyset$;
- (iii) Se $w \neq u$ con $|w| = |u|$, allora $O_w \cap K_u = \emptyset$;
- (iv) Se $\{w_0, \dots, w_{q-1}\}$ è una barriera stretta e consideriamo due suoi elementi $w_a \neq w_b$, allora $O_{w_a} \cap O_{w_b} = \emptyset$.

Dimostrazione. (i) Ovvio.

(ii) Sia $x \in K$, $x = \pi(\mathbf{i})$ e $y \in O$ qualunque. Allora dal teorema fondamentale sugli IFS, dato che $\{y\}$ è compatto si ha che $S_{\mathbf{i}|n}(y) \xrightarrow{n \rightarrow \infty} x$. Dato che Φ manda O in sé si ha che per ogni n , $S_{\mathbf{i}|n}(y) \in O$ per cui x è un punto di accumulazione per O , e quindi

$K \subseteq O^f$. La tesi segue applicando S_w ad entrambi i membri dell'inclusione, dato che questa operazione commuta con quella di chiusura.

(iii) Siano $w \neq u$ con $|w| = |u|$. Allora per ipotesi $O_w \cap O_u = \emptyset$. Poiché sono entrambi aperti, vale anche $O_w \cap O_u^f = \emptyset$, e quindi poiché per (ii) si ha $K_u \subseteq O_u^f$, vale $O_w \cap K_u = \emptyset$.

(iv) Presi w e u come nell'enunciato, essendo parte di una barriera stretta questi sono inconfrontabili. Ma allora $w = vi \dots$ e $u = vj \dots$ per un qualche v eventualmente vuoto, con $i \neq j$ la prima lettera su cui differiscono. Proprio poiché $i \neq j$, $O_i \cap O_j = \emptyset$ per cui applicando S_v ad entrambi i membri otteniamo $O_{vi} \cap O_{vj} = \emptyset$, da cui per il punto (i) $O_w \cap O_u = \emptyset$. \square

s-insiemi

Sappiamo esibire molti esempi di s -insiemi. Meno ovvio è trovare esempi di oggetti che *non* sono s -insiemi.

Esempio 5 (Successione di insiemi di Cantor). Consideriamo una successione $0 < r_1 < r_2 < \dots$ con $r_n \rightarrow 1/2$, e l'IFS associato composto dalle due mappe che danno l'insieme di Cantor opportunamente corrette: $x \mapsto r_n x$, $x \mapsto r_n x + (1 - r_n)$. Esso ha un certo attrattore C_n che ha misura di Lebesgue 0. D'altro canto però vale banalmente la OSC con $O = (0, 1)$, per cui $\dim_H C_n = D_n$ con D_n dato da $2r_n^{D_n} = 1$, che ha come soluzione

$$D_n = \frac{\log(1/2)}{\log r_n} \xrightarrow{n \rightarrow \infty} 1.$$

Questo implica che $\bigcup_{n \geq 0} C_n$ ha ancora misura di Lebesgue 0 (è unione numerabile di insiemi di λ -misura nulla) ma dimensione di Hausdorff 1. Non è dunque un 1-insieme.

Esempio 6 (Numeri malapprossimabili).

2.13 Densità locale e similitudini

C'è un interessante risultato che riguarda la misura stazionaria nel caso di IFS di similitudini. Per arrivarci ci occorre enunciare due teoremi importanti e dimostrare un lemma.

Definizione 36. Fissato $x \in \mathbb{R}^n$, $\mu \in \mathcal{M}$ e $s \in \mathbb{R}_{\geq 0}$, la *densità locale s -dimensionale inferiore [superiore] di μ in x* è

$$\underline{\theta}^s(\mu, x) := \liminf_{\varepsilon \rightarrow 0} \frac{\mu(B(x, \varepsilon))}{\gamma(s) \cdot \varepsilon^s}$$

$$\left[\bar{\theta}^s(\mu, x) := \limsup_{\varepsilon \rightarrow 0} \frac{\mu(B(x, \varepsilon))}{\gamma(s) \cdot \varepsilon^s} \right]$$

dove con $B(x, \varepsilon)$ intendiamo la palla chiusa di centro x e raggio ε , con $\gamma(s) = \frac{\pi^{s/2}}{\Gamma(s/2+1)}$ la misura della palla unitaria s -dimensionale.

Quindi in pratica la densità locale misura quanto è “massiva” una palla infinitesimale centrata in x misurata con μ rispetto a quella misurata con la misura di Lebesgue.

Teorema 26 (di densità di Lebesgue). *Sia $A \subseteq \mathbb{R}^n$ Lebesgue-misurabile con $\lambda(A) > 0$. Allora per λ -ogni $x \in \mathbb{R}^n$ si ha*

$$\underline{\theta}^n(\lambda|_A, x) = \bar{\theta}^n(\lambda|_A, x) = \begin{cases} 0 & x \notin A \\ 1 & x \in A. \end{cases}$$

Si ha una parziale inversa di tale teorema.

Teorema 27 (Federer). *Sia $\mu \in \mathcal{M}$ una misura finita e A un boreliano. Allora:*

1. *se per una certa costante d_1 per cui $\forall x \in A$ $d_1 \leq \bar{\theta}^s(\mu, x)$, allora $d_1 \cdot \mathcal{H}^s(A) \leq \mu(A)$;*
2. *se per una costante d_2 vale $\forall x \in A$ $\bar{\theta}^s(\mu, x) \leq d_2$, allora $\mu(A) \leq 2^s d_2 \cdot \mathcal{H}^s(A)$.*

In particolare, se $\bar{\theta}^s(\mu, x)$ è limitata sia dall'alto che dal basso in modo indipendente da x allora

$$\mu(A) \neq 0, +\infty \iff \mathcal{H}^s(A) \neq 0, +\infty.$$

Dimostriamo ora un lemma ed un teorema importanti.

Lemma 28. *Siano $0 < c_1 < c_2 < +\infty$, $x \in \mathbb{R}^n$, $r > 0$. Inoltre sia $\{O_i\}_{i \in I}$ una famiglia di aperti disgiunti tali che ogni O_i contiene una palla chiusa di raggio rc_1 ed è contenuto in una palla chiusa di raggio rc_2 . Allora*

$$\#\{i : B(x, r) \cap O_i^f \neq \emptyset\} \leq \left(\frac{1 + 2c_2}{c_1} \right)^n.$$

Dimostrazione. Supponiamo che O_0^f, \dots, O_{t-1}^f intersechino $B(x, r)$. Allora è facile vedere (provare a fare il disegno in \mathbb{R}^2) che $O_0^f, \dots, O_{t-1}^f \subseteq B(x, (1 + 2c_2)r)$. Poiché ognuno di questi aperti contiene una pallina chiusa $B(a_i, rc_1)$, si ha in particolare

$$\bigcup_{i=0}^{t-1} B(a_i, rc_1) \subseteq B(x, (1 + 2c_2)r).$$

Ma allora passando alle misure vale

$$t\gamma(n)(c_1r)^n \leq \gamma(n)[(1 + 2c_2)r]^n,$$

da cui si trova $t \leq \left(\frac{1+2c_2}{c_1} \right)^n$. □

Teorema 29. *Sia $S = \{S_0, \dots, S_{N-1}\}$ un IFS di **similitudini** con fattori di contrazione $0 < r_0, \dots, r_{N-1} < 1$ con attrattore K e che soddisfa la OSC rispetto ad un aperto O . Sia D la dimensione di similitudine di S e poniamo $p_i := r_i^D$ per ogni i e $\mathbf{p} := (p_0, \dots, p_{N-1})$. Allora:*

1. $\exists c_1, c_2 : \forall x \in K, 0 < c_1 \leq \underline{\theta}^D(\mathcal{H}^D|_K, x) \leq \bar{\theta}^D(\mathcal{H}^D|_K, x) \leq c_2 < \infty$;
2. K è un D -insieme;
3. $\mu = \pi_* P_{\mathbf{p}} = \mathcal{H}^D(-|K) = \frac{\mathcal{H}^D|_K}{\mathcal{H}^D(K)}$.

Dimostrazione. Consideriamo inizialmente la misura stazionaria $\mu = \pi_* P$ e troviamo limitazioni dall'alto e dal basso per $\underline{\theta}^D(\mu, x) \leq \bar{\theta}^D(\mu, x)$ che siano indipendenti da x . Consideriamo un generico $x = \pi(\mathbf{i}) \in K$ ed $\varepsilon > 0$. Sia m il minimo indice per cui $K_{\mathbf{i}|m} \subseteq B(x, \varepsilon)$; allora $r_{\mathbf{i}|m} \cdot \text{diam } K \geq \varepsilon \cdot r_{\min}$, perché se così non fosse non sarebbe l' m minimo: infatti se valesse la disuguaglianza $r_{\mathbf{i}|m} \cdot \text{diam } K < \varepsilon \cdot r_{\min} \leq \varepsilon r_0, \dots, \varepsilon r_{N-1}$ avremmo $\text{diam } K_{\mathbf{i}|(m-1)} = r_{\mathbf{i}|(m-1)} \cdot \text{diam } K < \varepsilon$, assurdo.

Di conseguenza

$$\frac{\mu(B(x, \varepsilon))}{\gamma(D)\varepsilon^D} \geq \frac{\mu(K_{\mathbf{i}|m})}{\gamma(D)\varepsilon^D} = \frac{p_{\mathbf{i}|m}}{\gamma(D)\varepsilon^D} = \frac{1}{\gamma(D)} \left(\frac{r_{\mathbf{i}|m}}{\varepsilon} \right)^D \geq \frac{1}{\gamma(D)} \left(\frac{r_{\min}}{\text{diam } K} \right)^D$$

quindi $\underline{\theta}^D(\mu, x) \geq \frac{1}{\gamma(D)} \left(\frac{r_{\min}}{\text{diam } K} \right)^D > 0$.

Fissiamo ora $0 < c_1 < c_2 < \infty$ in modo che l'insieme O dato per ipotesi contenga una palla chiusa di raggio c_1 e sia contenuto in una palla chiusa di raggio c_2 . Per ogni $\mathbf{j} \in N^\omega$, sia $q \geq 0$ minimo tale che

$$r_{\min} \varepsilon \stackrel{(*)}{\leq} r_{\mathbf{j}|q} \stackrel{(**)}{\leq} \varepsilon.$$

Allora la famiglia di tutti i blocchi del tipo $[\mathbf{j} \upharpoonright q]$ copre N^ω per cui dato che N^ω è compatto possiamo estrarne una sottosuccessione finita (cioè una barriera stretta), sia questa $\mathcal{B} = \{w_0, \dots, w_{t-1}\}$. Per un vecchio lemma si ha che $O_{w_0}, \dots, O_{w_{t-1}}$ sono aperti disgiunti. Inoltre ogni O_w con $w \in \mathcal{B}$ contiene una palla di raggio $r_w c_1$ perché O ne contiene una di raggio c_1 . Dunque, per la disuguaglianza $(*)$, O_w contiene anche una palla chiusa di raggio $r_{\min} \varepsilon c_1$, e inoltre è contenuto in una palla chiusa di raggio $r_w c_2$, per cui per $(**)$ in una di raggio εc_2 . A questo punto possiamo applicare il lemma precedente, ottenendo

$$\#\{w \in \mathcal{B} : B(x, \varepsilon) \cap O_w^f \neq \emptyset\} \leq \left(\frac{1 + 2c_2}{r_{\min} c_1} \right)^n.$$

Ora, poiché $\forall u \in N^*$ si ha $K_u \subseteq O_u^f$, la stessa limitazione vale anche per la cardinalità dell'insieme

$$\mathcal{B}' := \{w \in \mathcal{B} : B(x, \varepsilon) \cap K_w \neq \emptyset\}.$$

Osservando che $\text{supp } \mu \subseteq K$ e $K = \cup_{w \in \mathcal{B}} K_w$, abbiamo

$$\frac{\mu(B(x, \varepsilon))}{\gamma(D)\varepsilon^D} \leq \frac{1}{\gamma(D)\varepsilon^D} \sum_{w \in \mathcal{B}'} \mu(K_w) \leq \frac{1}{\gamma(D)\varepsilon^D} \left(\frac{1 + 2c_2}{r_{\min} c_1} \right)^n \varepsilon^D$$

dove l'ultima disuguaglianza segue dal fatto che $\mu(K_w) \leq \varepsilon^D$ perché se $w \in \mathcal{B}' \subseteq \mathcal{B}$, si ha $\mu(K_w) = P([w]) = p_w = r_w^D \leq \varepsilon^D$ di nuovo per $(**)$. Dalla disuguaglianza appena trovata deduciamo quindi che

$$\bar{\theta}^D(\mu, x) \leq \frac{1}{\gamma(D)} \left(\frac{1 + 2c_2}{r_{\min} c_1} \right)^n < \infty.$$

Abbiamo dunque determinato due costanti m_1, m_2 per cui vale $0 < m_1 \leq \underline{\theta}^D(\mu, x) \leq \bar{\theta}^D(\mu, x) \leq m_2 < \infty$. Applicando il teorema di Federer abbiamo che $\mu(K) \neq 0, +\infty$ se e solo se $\mathcal{H}^D(K) \neq 0, +\infty$. Ma $\mu(K) = 1$, per cui K risulta un D -insieme ed abbiamo dimostrato il punto 2.

L'uguaglianza da dimostrare del punto (3) è la seconda, e lo facciamo riciclando una vecchia idea: se dimostriamo che il termine di destra è una misura stazionaria, allora deve essere l'unica misura stazionaria che sappiamo già essere proprio $\mu = \pi_* P$. Ricordando un lemma visto in precedenza, $s = D \iff \forall i \neq j, \mathcal{H}^s(K_i \cap K_j) = 0$. Ne segue che

$$\mathcal{H}^D \llcorner K = \sum_i \mathcal{H}^D \llcorner K_i \stackrel{(!)}{=} \sum_i r_i^D(S_i)_*(\mathcal{H}^D \llcorner K) = \Phi(\mathcal{H}^D \llcorner K)$$

da cui per l'unicità (a meno di costanti moltiplicative) del punto fisso, segue

$$\mu = \frac{\mathcal{H}^D \llcorner K}{\mathcal{H}^D(K)} = \mathcal{H}^D(-|K).$$

Verifichiamo l'uguaglianza (!), dimostrando che se S è una similitudine di fattore r e $s \geq 0$, allora vale

$$\mathcal{H}^s \llcorner (SK) = r^s(S_*)(\mathcal{H}^s \llcorner K).$$

Infatti

$$\begin{aligned} [\mathcal{H}^s \llcorner (SK)](A) &= \mathcal{H}^s(SK \cap A) = \mathcal{H}^s(S(K \cap S^{-1}A)) = \\ &= r^s \mathcal{H}^s(K \cap S^{-1}A) = r^s(\mathcal{H}^s \llcorner K)(S^{-1}A) = r^s[S_*(\mathcal{H}^s \llcorner K)](A) \end{aligned}$$

per cui anche il punto 3 è stato dimostrato.

Infine, nella prima parte della dimostrazione avevamo ottenuto le limitazioni

$$0 < m_1 \leq \underline{\theta}^D(\mu, x) \leq \bar{\theta}^D(\mu, x) \leq m_2 < \infty$$

e ora sappiamo anche che $\mu = \mathcal{H}^D(-|K)$, per cui dato che le costanti moltiplicative passano fuori dalla densità superiore/inferiore, sostituendo si ha

$$0 < m_1 \leq \frac{\underline{\theta}^D(\mathcal{H}^D \llcorner K, x)}{\mathcal{H}^D(K)} \leq \frac{\bar{\theta}^D(\mathcal{H}^D \llcorner K, x)}{\mathcal{H}^D(K)} \leq m_2 < \infty.$$

Basta moltiplicare tutto per $\mathcal{H}^D(K)$ e si ottiene il punto 1. \square

2.14 Barnsley-Vince: IFS topologici, etc.

La nozione di IFS che abbiamo considerato finora si può generalizzare, definendo IFS un insieme di funzioni $f_n : X \rightarrow X$ continue, con X semplicemente spazio topologico. Tale insieme, che noi abbiamo sempre supposto finito, si potrebbe prendere infinito. Tuttavia noi manteniamo le ipotesi viste finora, ossia che X sia metrico e completo e che l'insieme di funzioni abbia cardinalità finita N .

Definizione 37. Un *attrattore* per l'IFS \mathcal{F} è un insieme $A \in \mathcal{K}(X)$ tale che

- $\Phi(A) = A$;
- esiste un aperto $U \subset X$ tale che $A \subset U$ e per ogni compatto $S \subset U$ vale il limite $\lim_{k \rightarrow \infty} \Phi^k(S) = A$, inteso nella metrica di Hausdorff.

Definizione 38. Il *bacino di attrazione* B di un attrattore A dell'IFS \mathcal{F} è il più grande insieme aperto U per cui vale la seconda condizione della definizione precedente.

Definizione 39. Il *repulsore duale* di un attrattore A dell'IFS \mathcal{F} è l'insieme $A^* := X \setminus B$, dove B è il bacino di attrazione di A .

Osservazione 40. Questi due insiemi dipendono strettamente dallo spazio X in cui si ambienta l'IFS: anche mantenendo lo stesso insieme di funzioni, bacino di attrazione e repulsore duale possono variare!

Definizione 41. L'IFS $\mathcal{F} = \{X; f_1, \dots, f_N\}$ si dice *invertibile* quando le mappe che lo compongono sono tutte degli omeomorfismi $X \rightarrow X$. In tal caso, l'IFS dato da $\mathcal{F}^{-1} := \{X; f_1^{-1}, \dots, f_N^{-1}\}$ si dice *IFS inverso* di \mathcal{F} .

Sotto alcune ipotesi, un attrattore per \mathcal{F}^{-1} è il repulsore duale di un attrattore di \mathcal{F} .

A seconda dello spazio ambiente X e del tipo di trasformazioni che compongono gli IFS, questi si classificano in modo intuitivo. Ad esempio se $X = \mathbb{R}^d$ e le funzioni sono affinità, si parla di IFS affine; se le funzioni sono similitudini, si parla di IFS di similitudini. Un caso che non abbiamo ancora affrontato è quello di $X = \mathbb{RP}^d$ con proiettività, che ovviamente si dirà IFS proiettivo. L'ultimo esempio che diamo è quello in cui $X = \hat{\mathbb{C}} = \mathbb{C} \cup \infty$ e le funzioni sono trasformazioni lineari fratte complesse, che si dicono IFS di Möbius.

Un fatto interessante è che la contrattività dell'IFS non è essenziale per l'esistenza di un attrattore nel caso degli IFS proiettivi: un controesempio è nell'articolo di Barnsley e Vince (2013). Valgono i seguenti risultati.

Teorema 30. *Un IFS affine, di Möbius o proiettivo può avere al più un attrattore.*

Teorema 31. 1. *Un IFS affine ha un attrattore se e solo se è contrattivo.*

2. *Un IFS di Möbius ha un attrattore $A \neq \hat{\mathbb{C}}$ se e solo se esiste un aperto non vuoto U tale che $U^f \neq \hat{\mathbb{C}}$ e l'IFS risulti contrattivo se ristretto a U .*

3. *Un IFS proiettivo ha un attrattore che non interseca almeno un iperpiano se e solo se esiste un aperto non vuoto U tale che l'IFS risulti contrattivo in U^f .*

Una precisazione che va fatta dopo aver enunciato questo risultato è che il fatto che l'IFS sia contrattivo non è necessariamente vero per tutte le possibili metriche che si possono considerare. Tuttavia, è sempre possibile definire un'opportuna metrica per cui un IFS con attrattore risulti contrattivo. Nel caso degli IFS affini tale metrica è quella di Minkowski, cioè

$$d_C(x, y) = \|x - y\|_C$$

dove si considera un corpo convesso centrato C che induce

$$\|x\|_C = \inf\{\lambda \geq 0 : x \in \lambda C\}.$$

Nel caso degli IFS proiettivi, la metrica giusta è una metrica di Hilbert modificata, definita sull'interno di un corpo convesso centrato C come

$$d_C(x, y) := \log R(a, b, x, y) = \log \frac{|ay||bx|}{|ax||by|}$$

dove a e b sono i punti in cui la retta che congiunge x e y interseca il bordo di C . Infine, un IFS di Möbius con attrattore è sempre contrattivo in un qualche insieme aperto U rispetto alla seguente metrica definita su $\hat{\mathbb{C}}$:

$$d_U(x, y) = \max_{z \notin U} \log \frac{|z - x|}{|z - y|} + \max_{x \notin U} \log \frac{|z - y|}{|z - x|}.$$

Ad ogni modo si può evitare il ricorso a metriche: infatti ogni IFS dei tre tipi citati che abbia un attrattore è anche contrattivo in un senso topologico.

Definizione 42. Un IFS \mathcal{F} si dice *topologicamente contrattivo* su un insieme compatto $K \subset X$ se $\Phi(K) \subset K^\circ$.

Teorema 32. 1. Un IFS affine ha un attrattore se e solo se Φ è topologicamente contrattivo su un qualche corpo convesso.

2. Un IFS di Möbius su $\hat{\mathbb{C}}$ ha un attrattore $A \neq \hat{\mathbb{C}}$ se e solo se Φ è topologicamente contrattivo su un qualche sottoinsieme proprio di $\hat{\mathbb{C}}$ compatto.

3. Un IFS proiettivo ha un attrattore che evita un iperpiano se e solo se Φ è topologicamente contrattivo sull'unione di un numero finito non nullo di corpi convessi disgiunti.

Infine citiamo un risultato che giustifica il *chaos game*, introducendo prima la nozione di spazio proprio.

Lemma 33. Sia (X, d) uno spazio metrico. Sono equivalenti:

1. i compatti sono tutti e soli i chiusi e limitati;
2. ogni palla chiusa è compatta.

Uno spazio che soddisfa tali condizioni si dice *proprio*.

Dimostrazione. L'implicazione (\Rightarrow) è ovvia. Vediamo l'altra. I compatti sono sempre chiusi e limitati. Viceversa, preso un chiuso C limitato da $R > 0$, se $a \in C$ allora il chiuso sta nella palla chiusa (e dunque compatta per ipotesi) $B(a, R)$. Tutti i chiusi in compatti sono compatti, quindi in particolare lo è anche C . \square

Esempio 7. \mathbb{R}^n e qualunque spazio X compatto sono propri. Un esempio di spazio improprio è ℓ_2 (in cui ad esempio la palla unitaria chiusa non è compatta, cfr. Istituzioni di Analisi Superiore).

Teorema 34. Sia X uno spazio metrico completo proprio e $\mathcal{F} = \{X; f_1, \dots, f_N\}$ un IFS con attrattore A di bacino B . Se $(x_n)_n$ è un'orbita casuale di $x_0 \in B$ tramite \mathcal{F} , allora tale orbita produce A quasi certamente, ossia

$$A = \lim_{K \rightarrow \infty} \{x_n : n \geq K\} = \bigcap_{K \geq 1} \overline{\{x_n : n \geq K\}}.$$

3 Joint spectral radius e norme estreme

3.1 Raggio spettrale congiunto

Esempio 8 (shearing del piano).

Il seguente lemma motiva una definizione che stiamo per dare.

Lemma 35. *Consideriamo una mappa lineare $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ senza sottospazi invarianti non banali (cioè senza autovalori reali). Allora esiste un'ellissi E e un numero reale $\lambda > 0$ tali che $L(E) = \lambda E$.*

Dimostrazione. Avendo L due autovalori complessi coniugati, esiste $S \in GL_2\mathbb{R}$ tale che

$$M := S^{-1}LS = \lambda \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

per un certo angolo θ e un $\lambda > 0$. Se consideriamo l'immagine $E = S(D)$ del disco unitario standard D , abbiamo facilmente da $M(D) = \lambda D$ che $L(E) = \lambda E$. \square

Definizione 43. Un IFS $\mathcal{R} = \{R_0, \dots, R_{N-1}\}$ con $R_i \in \mathcal{M}^{n \times n}\mathbb{R}$ si dice *irriducibile* se non esiste un sottospazio $V \subseteq \mathbb{R}^n$ non banale invariante in avanti per tutti gli R_i .

Notiamo che dall'irriducibilità non segue l'unicità dell'eventuale insieme invariante.

Esempio 9. Se consideriamo $\mathcal{R} = \left\{ \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \right\}$, abbiamo ovviamente un IFS irriducibile. Se $\theta/\pi \in \mathbb{Q}$, esistono un sacco (infiniti!) autoinsiemi di forme diverse. Se invece $\theta/\pi \notin \mathbb{Q}$, l'unico autoinsieme (sempre a meno di dilatazione) è il cerchio centrato nell'origine.

Iniziamo a parlare di raggio spettrale congiunto. Per poter dare la definizione occorrono prima dei risultati preliminari.

Dato $\mathcal{R} = \{R_0, \dots, R_{N-1}\}$ con $R_i \in \mathcal{M}^{n \times n}\mathbb{R}$, si pone

$$\mathcal{R}^n := \{R_w : w \in N^n\}, \quad \|\mathcal{R}^n\| := \max\{\|R_w\| : w \in N^n\}.$$

Lemma 36 (di Fekete). *Sia $a_1, a_2, \dots \in \mathbb{R}$ una successione subadditiva, ossia tale che $a_{n+m} \leq a_n + a_m$ per ogni n, m . Allora $\lim_{n \rightarrow \infty} \frac{a_n}{n}$ esiste e vale $\inf_{n \geq 1} \frac{a_n}{n}$. (Oss.: il limite può eventualmente valere $-\infty$).*

Dimostrazione. Poniamo $a := \inf_{n \geq 1} \frac{a_n}{n}$. Osservando che per subadditività

$$a_n \leq a_1 + a_{n-1} \leq 2a_1 + a_{n-2} \leq \dots \leq na_1,$$

abbiamo che $a \leq a_1$. Fissiamo ora $\varepsilon > 0$; per definizione di estremo inferiore esiste m con $a_m < m(a + \varepsilon)$. Facendo la divisione con resto, per ogni n vale

$$\begin{cases} n = q_n m + r_n \\ 0 \leq r_n < m. \end{cases}$$

Analogamente a quanto visto sopra, vale

$$a_n = a_{q_n m + r_n} \leq q_n a_m + a_{r_n} < q_n m(a + \varepsilon) + \underbrace{\sup\{0, a_1, \dots, a_{m-1}\}}_{=:b}.$$

Dunque

$$a \leq \liminf_{n \rightarrow \infty} \frac{a_n}{n} \leq \limsup_{n \rightarrow \infty} \frac{a_n}{n} \leq \limsup_{n \rightarrow \infty} \frac{q_n m(a + \varepsilon) + b}{q_n m + r_n} = a + \varepsilon$$

per cui per $\varepsilon \rightarrow 0$ tutto collassa su a e abbiamo concluso. \square

Corollario 37 (versione moltiplicativa). *Sia b_1, b_2, \dots sottomoltiplicativa ($b_{m+n} \leq b_m b_n$). Allora $\lim_{n \rightarrow \infty} b_n^{1/n}$ esiste e vale $\inf_n b_n^{1/n}$. Può eventualmente valere 0.*

Dimostrazione. Basta applicare exp al risultato del lemma di Fekete. \square

Applicando quest'ultimo corollario alla successione $b_1 = \|\mathcal{R}\|$, $b_2 = \|\mathcal{R}^2\|, \dots$, che è sottomoltiplicativa per le note proprietà delle norme operatoriali, possiamo giustificare la seguente

Definizione 44. Il numero $\tilde{\rho}(\mathcal{R}) = \lim_{n \rightarrow \infty} \|\mathcal{R}^n\|^{1/n} = \inf_n \|\mathcal{R}^n\|^{1/n}$ è detto *joint spectral radius* di \mathcal{R} . Sta in $\mathbb{R}_{\geq 0}$ ed è dominato da $\|\mathcal{R}\|$, quindi in realtà sta in $[0, \|\mathcal{R}\|]$.

Definizione 45. Una norma $\|\cdot\|$ si dice *estrema per \mathcal{R}* se $\tilde{\rho}(\mathcal{R}) = \|\mathcal{R}\|$, ossia se per essa il JSR risulta il massimo possibile.

Definizione 46. Preso $n \geq 1$, poniamo $\rho(\mathcal{R}^n) := \max\{\rho(R_w) : w \in N^n\}$, con ρ raggio spettrale in senso classico, ottenendo quindi una estensione di quest'ultimo ad un insieme finito di matrici.

Ci si potrebbe chiedere come sono legati tra loro il raggio spettrale “classico” e il joint spectral radius. Vale l'importante

Teorema 38 (di Berger-Wang). $\tilde{\rho}(\mathcal{R}) = \limsup_{n \rightarrow \infty} [\rho(\mathcal{R}^n)]^{1/n} = \sup_n [\rho(\mathcal{R}^n)]^{1/n}$.

Dimostrazione. Dimostriamo solo l'uguaglianza facile, la seconda. Senz'altro vale la disuguaglianza \leq , per cui resta da dimostrare l'altra. Per ogni n , se $r = \rho(\mathcal{R}^n)^{1/n}$ è realizzato da un certo $R_w \in \mathcal{R}^n$, allora $r \leq \rho(\mathcal{R}^{2n})^{1/(2n)}, \rho(\mathcal{R}^{3n})^{1/(3n)}, \dots$ in quanto il massimo raggio spettrale certamente non decresce (basti pensare che è assicurato che valga almeno r considerando R_w^2, R_w^3, \dots). Abbiamo quindi trovato tutta una sottosuccessione che vale almeno r , e di conseguenza vale anche la disuguaglianza \geq . \square

Osservazione 47. È noto che vale sempre $\rho(R) \leq \|R\|$, e di conseguenza passando ai massimi si ottiene $\rho(\mathcal{R}) \leq \|\mathcal{R}\|$. In particolare, il sup del teorema è dominato da $\|\mathcal{R}\| \geq \|\mathcal{R}^n\|^{1/n}$.

Esercizio 3. Si consideri l'insieme

$$\mathcal{R} = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$$

e ci si convinca che per definire il JSR non va bene usare il lim al posto del lim sup. (Suggerimento: fare un po' di moltiplicazioni tra matrici).

3.2 Corpi convessi e norme

Dopo aver definito il JSR, è ovviamente di notevole importanza trovare un modo per calcolarlo. A tale scopo introduciamo una interessante corrispondenza tra norme e “palle” in \mathbb{R}^n .

Definizione 48. Un insieme $C \subset \mathbb{R}^n$ si dice *corpo convesso centrato* (CCB) se è compatto, ha interno non vuoto ed è simmetrico rispetto allo $0 \in \mathbb{R}^n$.

Grazie al funzionale di Minkowski (talvolta chiamato anche *gauge*)

$$\|x\|_C = \inf\{\mu > 0 : x \in \mu C\}$$

si può stabilire una dualità $C \sim \|\cdot\|_C$ tra norme in \mathbb{R}^n e CCB. Vediamone alcune proprietà.

Proposizione 39. Dati C_1, C_2 CCB, allora:

- $C_1 \subseteq C_2 \iff \|\cdot\|_{C_1} \geq \|\cdot\|_{C_2}$;
- $C_1 \cap C_2$ è un CCB, e $\|\cdot\|_{C_1 \cap C_2} = \sup\{\|\cdot\|_{C_1}, \|\cdot\|_{C_2}\}$;
- se $t > 0$, allora $tC \sim t^{-1}\|\cdot\|_C$; più in generale, se $A \in GL_n \mathbb{R}$ allora $A[C] \sim \|A^{-1}\cdot\|_C$.

Dimostrazione. Il primo punto è ovvio. Per il secondo, si ha che preso $v \in \mathbb{R}^n$ vale

$$\begin{aligned} \|v\|_{C_1 \cap C_2} &= \min\{r \geq 0 : v \in r(C_1 \cap C_2)\} = \min\{r \geq 0 : v \in rC_1, v \in rC_2\} \\ &= \max\{\min\{r : v \in rC_1\}, \min\{r : v \in rC_2\}\} = \max\{\|v\|_{C_1}, \|v\|_{C_2}\}. \end{aligned}$$

Per l'ultimo punto, si vede che

$$\|v\|_{A[C]} = \min\{r \geq 0 : v \in rA[C]\} = \min\{r \geq 0 : A^{-1}v \in rC\} = \|A^{-1}v\|_C.$$

□

Osservazione 49. Dato un qualsiasi insieme $A \subseteq \mathbb{R}^n$ purché limitato e con interno non vuoto, esiste il minimo CCB che lo contiene, definito da $\bigcap\{C : C \text{ è un CCB}, A \subseteq C\}$. La condizione sull'interno non vuoto serve perché ad esempio si potrebbe pensare di fare l'intersezione infinita di rettangoli che si stringono su una retta, ottenendo un segmento che ovviamente non è un CCB.

Grazie all'osservazione, possiamo definire un sup sui corpi convessi come segue:

$$C_1 \vee C_2 := \bigcap\{C : C \text{ è un CCB}, C_1 \cup C_2 \subseteq C\}.$$

Una definizione alternativa dal basso è quella che consiste nel prendere $C_1 \cup C_2$ e farne la chiusura convessa. L'equivalenza con la definizione appena data è lasciata per esercizio. A livello di norme questa operazione di estremo superiore, se poniamo $C_3 = C_1 \vee C_2$, corrisponde alla norma definita da

$$\|v\|_{C_3} = \max\{\|v\| : \|v\| \text{ è una norma } \leq \text{ di } \|v\|_{C_1}, \|v\|_{C_2}\},$$

ossia all'estremo inferiore (fatto considerando l'ordinamento tra norme).

Ovviamente partendo da una norma indotta da un CCB si può definire come al solito la corrispondente norma di operatore, ricordando che C stesso è la palla unitaria per $\|\cdot\|_C$:

$$\|R\|_C = \|R[C]\|_C = \min\{r : R[C] \subseteq rC\}.$$

Esercizio 4 (carino). Si consideri la somma di Minkowski tra CCB definita da

$$C_1 + C_2 := \{c_1 + c_2 : c_1 \in C_1, c_2 \in C_2\}.$$

A cosa corrisponde per le norme?

Vediamo ora a cosa ci è servito introdurre la nozione di CCB e la dualità con le norme.

Lemma 40 (chiave). *Sia $w \in N^*$, e $[\rho(R_w)]^{1/|w|} = \alpha > 0$. Sia inoltre C un CCB tale che $\alpha^{-1}R_0C, \dots, \alpha^{-1}R_{N-1}C$ sono contenuti in C . Allora $\tilde{\rho}(\mathcal{R}) = \alpha$, la norma $\|\cdot\|_C$ è estrema per \mathcal{R} e R_w è un prodotto ottimo (cioè $\forall u \in N^*$, $[\rho(R_u)]^{1/|u|} \leq \alpha$). Inoltre necessariamente per almeno un indice $i \in N$ si ha $\alpha^{-1}R_iC \subsetneq C$.*

Dimostrazione. Vale senz'altro $\alpha \leq \tilde{\rho}$ per la caratterizzazione del JSR di Berger-Wang.

Sia ora $w = i_0i_1 \dots i_{q-1}$ una parola che soddisfa le ipotesi del teorema. Allora per ipotesi $R_{i_{q-1}}C \subseteq \alpha C$. Applicando $R_{i_{q-2}}$ ad entrambi i membri si ottiene quindi

$$R_{i_{q-2}}R_{i_{q-1}}C \subseteq \alpha R_{i_{q-2}}C \subseteq \alpha^2 C$$

e si ripete il procedimento fino ad arrivare a $R_wC \subseteq \alpha^{|w|}C$. Ne segue che $\|R_w\|_C \leq \alpha^{|w|}$ e siccome in generale $\tilde{\rho} \leq \|R_w\|_C^{1/|w|}$, abbiamo dimostrato anche l'altra disuguaglianza e quindi che $\tilde{\rho} = \alpha$.

Per ipotesi e definizione di norma operatoriale si ha che $\|\mathcal{R}\|_C \leq \alpha = \tilde{\rho}$, per cui dato che vale in generale $\tilde{\rho} \leq \|\mathcal{R}\|_C$ dato che è un estremo inferiore, vale l'uguaglianza $\|\mathcal{R}\|_C = \tilde{\rho}$ ossia la norma è estrema.

Il prodotto è ottimo perché nuovamente per Berger-Wang per ogni $u \in N^*$ si ha $\rho(R_u)^{1/|u|} \leq \tilde{\rho} = \alpha$.

Infine se fosse per assurdo $\alpha^{-1}R_iC \subsetneq C$ per ogni i , allora esisterebbe $0 < \beta < \alpha$ con $R_0C, \dots, R_{N-1}C \subseteq \beta C$. Ma allora $\|\mathcal{R}\|_C \leq \beta$, assurdo perché questa norma è estrema per \mathcal{R} .

Dimostrazione alternativa (da verificare): per Berger-Wang vale $\alpha \leq \tilde{\rho}$, in generale $\tilde{\rho} \leq \|\mathcal{R}\|_C$ e per l'ipotesi sui contenimenti vale infine $\|\mathcal{R}\|_C \leq \alpha$. Di conseguenza mettendo tutto in fila si trova

$$\alpha \leq \tilde{\rho} \leq \|\mathcal{R}\|_C \leq \alpha.$$

Poiché il primo e l'ultimo termine sono uguali, anche le disuguaglianze sono in realtà tutte uguaglianze, e questo dimostra sia che $\tilde{\rho} = \alpha$ sia che la norma è estrema. \square

Il lemma chiave ci permette di definire una “procedura universale” per calcolare il JSR, che si riduce a questo punto a trovare una norma ottima (o equivalentemente il CCB che la induce). Tale procedura consiste nei seguenti passi:

- congetturare una parola ottima w ;
- calcolare l'autovettore dominante v risolvendo $R_w v = \alpha^{|w|} v$;
- fissata una lunghezza l_{\max} , calcolare l'involuppo convesso di $\{\alpha^{-|u|} R_u v : |u| \leq l_{\max}\}$ e prenderlo come C .

Spesso tale procedura restituisce una norma estrema, e dunque il JSR.

Esempio 10 (*shearing* del piano).

3.3 JSR e autoinsiemi

Lemma 41. *Se $\tilde{\rho}(\mathcal{R}) < 1$ allora esiste C CCB tale che $\mathcal{R}(C) \subseteq \overset{\circ}{C}$.*

Dimostrazione. Sia $\|\cdot\|$ una norma corrispondente ad un certo CCB D . Poiché $\tilde{\rho} < 1$, esiste $m > 1$ tale che $\|\mathcal{R}^m\|^{1/m} < 1$. Dunque $\|\mathcal{R}^m\| < 1$ e $\mathcal{R}^m(D) \subseteq \overset{\circ}{D}$. Di conseguenza anche $\text{conv}(\mathcal{R}^m(D)) \subseteq \overset{\circ}{D}$. Verifichiamo che l'insieme $C := \sum_{k=0}^{m-1} \text{conv}(\mathcal{R}^k(D))$ è quello cercato. Infatti per ogni $R \in \mathcal{R}$ si ha

$$\begin{aligned} R[C] &= \sum_{k=0}^{m-1} \text{conv}(R\mathcal{R}^k(D)) \subseteq \sum_{k=0}^{m-1} \text{conv}(\mathcal{R}^{k+1}(D)) \\ &= \sum_{k=1}^m \text{conv}(\mathcal{R}^k(D)) \subseteq \overset{\circ}{D} + \sum_{k=1}^{m-1} \text{conv}(\mathcal{R}^k(D)) = \overset{\circ}{C} \end{aligned}$$

dove l'ultima uguaglianza segue perché (esercizio) per ogni coppia di corpi convessi K_1, K_2 vale $\overset{\circ}{K}_1 + K_2 = (\overset{\circ}{K}_1 + \overset{\circ}{K}_2)$, e inoltre $D = \text{conv} \mathcal{R}^0(D)$ perché è già convesso e $\mathcal{R}^0 = \{\text{Id}\}$ per convenzione. \square

Teorema 42. *Se \mathcal{R} è irriducibile e vale $\tilde{\rho} = \tilde{\rho}(\mathcal{R}) > 0$, allora esiste una norma estrema per \mathcal{R} .*

Dimostrazione. Possiamo sempre riscalarare \mathcal{R} dividendo per $\tilde{\rho}$, per cui senza perdita di generalità possiamo supporre che $\tilde{\rho} = 1$.

Ora scegliamo una successione $0 < \lambda_1 < \lambda_2 < \dots \rightarrow 1$, e osserviamo che $\tilde{\rho}(\lambda_k \mathcal{R}) = \lambda_k < 1$. Per il lemma precedente, per ogni k esiste un C_k CCB tale che $(\lambda_k \mathcal{R})(C_k) \subseteq \overset{\circ}{C}_k$. Sempre senza perdita di generalità possiamo riscalarare tutti i C_k in modo tale che tocchino la palla unitaria standard B sul bordo. Poiché siamo in dimensione finita B è compatta e di conseguenza $\mathcal{K}(B)$ è compatto con la metrica di Hausdorff e quindi anche sequenzialmente compatto. In particolare, possiamo trovare una sottosuccessione di compatti $C_{k_j} \xrightarrow{j \rightarrow \infty} C \in \mathcal{K}(B)$ con la convergenza pensata nella metrica di Hausdorff.

Vediamo che C è effettivamente un CCB con interno non vuoto (ossia *assorbente*). È facile verificare che se i C_{k_j} sono convessi e centrati allora lo è anche C (basta applicare la definizione di convessità e quella di d_H). Inoltre C non è ridotto al solo 0 perché per convergenza nella metrica di Hausdorff ogni suo ε -gonfiaggio dovrebbe inglobare quasi tutti i C_{k_j} ma questi toccano ∂B , assurdo. Ora bisogna vedere che C è assorbente. Poiché \mathcal{R} è irriducibile, gli unici sottospazi invarianti possono essere $\{0\}, \mathbb{R}^n$. Il primo lo è sempre, il secondo lo sarà appena dimostriamo che $\mathcal{R}(C) \subseteq C$, perché in tal caso lo span lineare di C è un sottospazio invariante non ridotto al solo 0.

A questo punto resta da verificare che la norma indotta da C è estrema, ossia che per ogni i vale $R_i[C] \subseteq C$. Fissiamo $R = R_i$ e osserviamo che per ogni j si ha, per come abbiamo definito i C_k , che $(\lambda_{k_j} R)(C_{k_j}) \subseteq \overset{\circ}{C}_{k_j}$. Vediamo che ciò implica $R_i[C] \subseteq C$. Infatti, sia $c \in C$; allora possiamo trovare una successione di $c_{k_j} \in C_{k_j}$ con $c_{k_j} \xrightarrow{j \rightarrow \infty} c$. Poiché R è continua, pure $Rc_{k_j} \xrightarrow{j \rightarrow \infty} Rc$. D'altronde $Rc_{k_j} \in \lambda_{k_j}^{-1} \overset{\circ}{C}_{k_j} \supseteq C_{k_j}$ e si può trovare un certo $c'_{k_j} \in C_{k_j}$ tale che $|c'_{k_j} - Rc_{k_j}| < 1 - \lambda_{k_j}$. Si conclude con la disuguaglianza

triangolare che $c'_{k_j} \rightarrow Rc$, e quindi $Rc \in C$ perché se non ci stessee avremmo un certo ε -gonfiaggio di C che esclude quasi tutti i punti della successione c'_{k_j} , assurdo. \square

Teorema 43. *Sia \mathcal{R} irriducibile con $\tilde{\rho} > 0$. Allora esiste un $\tilde{\rho}$ -autoinsieme E , ossia tale che $\mathcal{R}(E) = \tilde{\rho}E$, E è compatto, $[-1, 1]E = E$ e il \mathbb{R} -span è tutto \mathbb{R}^n .*

Dimostrazione. Senza perdita di generalità $\tilde{\rho} = 1$. Sia C il CCB estremo dato dal teorema precedente. Allora

$$C \supseteq \mathcal{R}(C) \supseteq \mathcal{R}^2(C) \supseteq \mathcal{R}^3(C) \supseteq \dots$$

Poniamo $E := \bigcap_{k \geq 0} \mathcal{R}^k(C)$.

Ovviamente E è compatto e \mathcal{R} -invariante. Ogni $\mathcal{R}^k C$ è $[-1, 1]$ -invariante (per induzione, dato che lo è C), quindi lo è anche E .

Se dimostriamo che $E \neq \{0\}$ allora il suo span lineare è tutto \mathbb{R}^n per l'ipotesi di irriducibilità. Per ogni $k \geq 1$ si ha che $\mathcal{R}^k(C) \subsetneq \overset{\circ}{C}$ perché $\tilde{\rho} = 1$. Quindi $\exists v_k \in \mathcal{R}^k(C) \cap \partial C$. Poiché ∂C è compatto, si può estrarre una sottosuccessione in esso contenuta tale che $v_{k_j} \rightarrow v \in \partial C$. Se per assurdo fosse $v \notin E$, esisterebbe $\varepsilon > 0$ tale che, posto $U(v, \varepsilon)$ il disco aperto di centro v e raggio ε , avremmo $U(v, \varepsilon) \cap \mathcal{R}^{k_j}(C) = \emptyset$ per ogni j sufficientemente grande, che è assurdo per l'esistenza della successione v_k di prima. \square

Gli ultimi due risultati visti assumono una notevole importanza alla luce del seguente Fatto: se \mathcal{R} è irriducibile, allora $\tilde{\rho}(\mathcal{R}) > 0$. Questo viene dal

Teorema 44 (di Levitzki). *Se un semigruppato di matrici ha tutti gli elementi nilpotenti, allora c'è un sottospazio invariante non banale.*

Dimostrazione del Fatto. Poiché in $\langle \mathcal{R} \rangle$ non ci sono sottospazi invarianti non banali per l'ipotesi di irriducibilità, per il Teorema di Levitzki esiste $k \geq 1$ e $R_w \in \mathcal{R}^k$ che non è nilpotente. Poiché questo è equivalente a dire che $\rho(R_w) > 0$, e per la caratterizzazione di Berger-Wang $\tilde{\rho} = \sup\{\rho(\mathcal{R}^k)^{1/k} : k \geq 1\}$, allora anche $\tilde{\rho} > 0$. \square

3.4 Norme di Protasov e Barabanov

Definizione 50. Dato \mathcal{R} irriducibile, un *norma di Protasov* è una norma $\|\cdot\| \sim C$ che soddisfa $C = \bigvee_{R \in \mathcal{R}} \tilde{\rho}^{-1} R[C]$. Dualmente, una *norma di Barabanov* è una norma che soddisfa $\|\cdot\| = \bigvee_{R \in \mathcal{R}} \tilde{\rho}^{-1} \|R \cdot -\|$.

Lemma 45. *Valgono i seguenti risultati:*

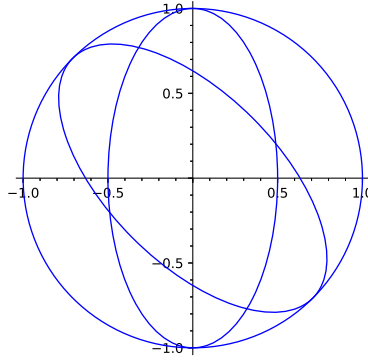
1. *Sia le norme di Protasov che quelle di Barabanov sono estreme;*
2. *Esistono norme estreme che non sono né di Protasov né di Barabanov;*
3. *Supponiamo che $\mathcal{R} \subseteq GL_n \mathbb{R}$. Allora $\|\cdot\|$ è di Protasov se e solo se soddisfa $\|\cdot\| = \bigwedge_{R \in \mathcal{R}} \tilde{\rho} \|R^{-1} \cdot -\|$;*
4. *C è di Barabanov se e solo se $C = \bigcap_{R \in \mathcal{R}} \tilde{\rho} R^{-1} C$ se e solo se C è estrema e inoltre $\forall v \in \partial C$ esiste $R \in \mathcal{R}$ tale che $\tilde{\rho}^{-1} Rv \in \partial C$.*

Dimostrazione. (1) Ricordiamo che estrema vuol dire $\tilde{\rho}(\mathcal{R}) = \|\mathcal{R}\|$ e che \leq vale sempre. A questo punto per le norme di Protasov è ovvio, quindi lo dimostriamo per quelle di Barabanov. Fissati $R \in \mathcal{R}$ e $c \in C$, abbiamo $1 \geq \|c\| \geq \tilde{\rho}^{-1}\|Rc\|$, per cui $\tilde{\rho}^{-1}Rc \in C$.

(2) Diamo un esempio esplicito considerando

$$R_0 = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 3/4 & -1/4 \\ -1/4 & 3/4 \end{pmatrix}$$

che sono rispettivamente una contrazione lungo l'asse x di fattore $1/2$ e la stessa a cui viene composta una rotazione di $\pi/4$. L'immagine della circonferenza unitaria tramite le due mappe si può vedere in figura. Abbiamo quindi che per $\mathcal{R} = \{R_0, R_1\}$ la norma



euclidea standard è estrema, ma non è di Protasov perché ovviamente l'involuppo convesso delle due ellissi non dà tutto il cerchio e non è nemmeno di Barabanov (si consideri ad esempio il punto $(1, 0)$).

Il punto (3) segue dal formalismo della dualità, così come la prima equivalenza del punto (4). Dimostriamo l'ultima equivalenza. (\Rightarrow) che la norma sia estrema segue dal punto (1). Sia ora $v \in \partial C$. Allora $\|v\| = 1$ e per definizione di norma di Barabanov esiste $R \in \mathcal{R}$ con $\tilde{\rho}^{-1}\|Rv\| = 1$, ma questo vuol dire che $\tilde{\rho}^{-1}Rv \in \partial C$. (\Leftarrow) Sia ora v qualsiasi. Senza perdita di generalità, a meno di normalizzarlo possiamo supporre che valga $\|v\| = 1$. Vogliamo quindi dimostrare che $1 = \bigvee_{R \in \mathcal{R}} \tilde{\rho}^{-1}\|Rv\|$. Usando l'estremalità della norma abbiamo che

$$\forall R \in \mathcal{R}, \quad \tilde{\rho}^{-1}\|Rv\| \leq \tilde{\rho}^{-1}\|R\| \underbrace{\|v\|}_{=1} \leq \tilde{\rho}^{-1}\tilde{\rho} = 1.$$

Per la disuguaglianza opposta, basta osservare che per la R data dalla condizione si ha $\tilde{\rho}^{-1}\|Rv\| = 1$, quindi il sup deve essere almeno tale valore. \square

Osservazione 51. Nel punto (4) non si chiede che R sia invertibile. In effetti quanto detto funziona con le seminorme, ossia quelle per cui $\|v\| = 0$ non implica $v = 0$. L'unica differenza è che in luogo dei CCB assorbenti si usano dei chiusi centrati assorbenti. Ad esempio si può considerare come “corpo” la striscia infinita $\mathbb{R} \times [-1, 1]$. Esso induce la seminorma $\|(x, y)\| = |y|$. A questo punto considerando come $R = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ e come C la palla unitaria standard (di \mathbb{R} , lungo l'asse y !), $R^{-1}C$ è proprio la striscia considerata.

Diamo ora un corollario ad un teorema della sezione precedente.

Corollario 46. *Sia \mathcal{R} irriducibile. Allora esiste una norma di Protasov per \mathcal{R} .*

Dimostrazione. A meno di dividere per $\tilde{\rho}$ possiamo supporre senza perdita di generalità che $\tilde{\rho} = 1$. Sia E l'autoinsieme dato dal teorema, vediamo che $C = \text{conv}(E)$ è di Protasov. Ovviamente è compatto, convesso e simmetrico. Inoltre, poiché lo span lineare di E è tutto \mathbb{R}^n , C è pure assorbente, quindi definisce una norma tramite il funzionale di Minkowski. Bisogna verificare che è di Protasov, ossia che $C = \text{conv}(\mathcal{R}(C))$. Si ha che

$$\begin{aligned} \text{conv} \left(\bigcup_{R \in \mathcal{R}} R[C] \right) &= \text{conv} \left(\bigcup_{R \in \mathcal{R}} R[\text{conv } E] \right) = \text{conv} \left(\bigcup_{R \in \mathcal{R}} \text{conv}(RE) \right) \stackrel{*}{=} \\ &= \text{conv} \left(\bigcup_{R \in \mathcal{R}} RE \right) = \text{conv}(\mathcal{R}[E]) = \text{conv}(E) = C. \end{aligned}$$

In $*$ basta osservare che un convesso contiene l'unione $A_1 \cup \dots \cup A_t$ se e solo se contiene $\text{conv } A_1 \cup \dots \cup \text{conv } A_t$, e nel passaggio seguente si usa il fatto che E sia un autoinsieme. \square

Analizziamo ora cosa succede considerando lo spazio duale $(\mathbb{R}^n)^*$ identificato con \mathbb{R}^n tramite $x \mapsto \langle x, - \rangle$ (attenzione: la x di partenza è il funzionale, quella di arrivo il vettore che lo rappresenta). Un elemento del duale è un funzionale lineare $x : (\mathbb{R}^n, \|\cdot\|) \rightarrow (\mathbb{R}, |\cdot|)$, per cui la sua norma operatoriale, fissato un CCB C , è data da

$$\|x\|^* = \max_{v \neq 0} \frac{|\langle x, v \rangle|}{\|v\|} = \max_{\|v\| \leq 1} |\langle x, v \rangle| = \max \langle x, C \rangle$$

dove nell'ultima espressione possiamo togliere il valore assoluto in quanto C è simmetrico. Ha quindi senso definire il duale di un CCB come

$$C^* = \{x : \|x\|^* = 1\} = \{x : \langle x, C \rangle \stackrel{(\equiv)}{\leq} 1\}.$$

Vediamo ora che questa è una “vera” dualità.

Proposizione 47. $C^{**} = C$.

Dimostrazione. Scriviamo per bene cos'è il biduale di un corpo convesso:

$$C^{**} = \{v : \langle C^*, v \rangle \leq 1\} = \{v : \forall x \in C^*, \langle x, v \rangle \leq 1\} = \{v : \forall x (\langle x, C \rangle \leq 1 \implies \langle x, v \rangle \leq 1)\}.$$

A questo punto (\supseteq) è ovvio perché se $v \in C$ l'implicazione tra parentesi è automaticamente vera, per cui vediamo (\subseteq) . Ci occorre innanzitutto fare l'osservazione chiave che ogni CCB in \mathbb{R}^n è l'intersezione di tutti i semispazi chiusi che lo contengono. Questo segue dal teorema di Hahn-Banach, che afferma che se $g : \mathbb{R}^n \rightarrow \mathbb{R} \geq 0$ è subadditiva e W è un sottospazio e $f : W \rightarrow \mathbb{R}$ è lineare con $|f| \leq g$, allora esiste \tilde{f} che estende f a tutto \mathbb{R}^n ed è ancora dominata da g .

Nel nostro caso andiamo a dimostrare che $v \notin C \implies v \notin C^{**}$. Preso v , consideriamo il sottospazio $W := \langle v \rangle$ e su di esso definiamo

$$f(sv) := s\|v\|, \quad s \in \mathbb{R}$$

che è lineare, subadditiva e dominata: $|f(sv)| \leq \|sv\|$. Abbiamo allora che per H-B esiste \tilde{f} che estende f , e che possiamo pensare come $\tilde{f} = \langle x, - \rangle$ per un qualche x , ancora dominata: $|\langle x, - \rangle| \leq \| - \|$. Tale x testimonia che $v \notin C^{**}$, infatti se $c \in C$ allora $|\langle x, c \rangle| \leq \|c\| \leq 1$ ma $|\langle x, v \rangle| = |f(v)| = \|v\| > 1$ in quanto $v \notin C$. \square

Dato un IFS \mathcal{R} , ad esso si associa $\mathcal{R}^T = \{R^T : R \in \mathcal{R}\}$. Valgono allora alcune considerazioni riassunte nel seguente enunciato.

Lemma 48. 1. \mathcal{R} è irriducibile se e solo se \mathcal{R}^T è irriducibile;

$$2. R[C] \subseteq C \iff R^T[C^*] \subseteq C^*;$$

$$3. (R[C])^* = (R^T)^{-1}[C^*] \text{ (anche se } R \text{ non è invertibile);}$$

$$4. (\bigvee_i C_i)^* = \bigcap_i C_i^*.$$

Dimostrazione. (1) Sia $0 < W < \mathbb{R}^n$ un sottospazio \mathcal{R} -invariante. Allora W^\perp è proprio (ovvio per dimensionalità) e \mathcal{R}^T -invariante. Infatti presi $R \in \mathcal{R}$ e $v \in W^\perp$, si ha che

$$\forall w \in W, \langle R^T v, w \rangle = \langle v, Rw \rangle = 0$$

in quanto $Rw \in W$ per ipotesi. L'altra implicazione si fa scambiando il ruolo di \mathcal{R} e \mathcal{R}^T . (2) è ovvio. (3) segue dalla seguente catena di equivalenze:

$$x \in (R[C])^* \iff \forall c \in C, \langle x, Rc \rangle \leq 1 \iff \forall c \in C, \langle R^T x, c \rangle \leq 1 \iff R^T x \in C^*.$$

(4) segue da

$$x \in \left(\bigvee_i C_i \right)^* \iff \forall v \in \text{conv} \left(\bigcup_i C_i \right), \langle x, v \rangle \leq 1 \iff \forall v \in \bigcup_i C_i, \langle x, v \rangle \leq 1.$$

\square

La dualità raggiunge il suo culmine nel seguente

Lemma 49 (di Wirth). *Sia \mathcal{R} irriducibile. Allora C è di Protasov [Barabanov] per \mathcal{R} se e solo se C^* è di Barabanov [Protasov] per \mathcal{R}^T .*

Dimostrazione. Il caso tra parentesi segue dall'altro gratis poiché $C^{**} = C$ e $(R^T)^T = R$.

Osserviamo che $\tilde{\rho}(\mathcal{R}) = \tilde{\rho}(\mathcal{R}^T)$ e dunque senza perdita di generalità a meno di dividere possiamo supporre $\tilde{\rho} = 1$. Allora C è di Protasov per \mathcal{R} se e solo se (lemma) $C = \bigvee_{R \in \mathcal{R}} R[C]$ se e solo se (dualità)

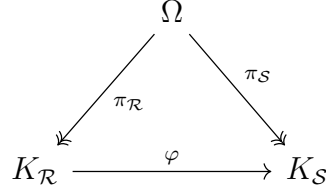
$$C^* = \left(\bigvee_{R \in \mathcal{R}} R[C] \right)^* = \bigcap_{R \in \mathcal{R}} (R[C])^* = \bigcap_{R \in \mathcal{R}} (R^T)^{-1}(C^*) = \bigcap_{R \in \mathcal{R}^T} R^{-1}(C^*)$$

se e solo se (lemma) C^* è di Barabanov per \mathcal{R}^T . \square

4 Trasformazioni frattali

Siano dati $\mathcal{R} = \{R_0, \dots, R_{N-1}\}$ e $\mathcal{S} = \{S_0, \dots, S_{N-1}\}$ non necessariamente strettamente contrattivi, però tali per cui il teorema fondamentale degli IFS sia valido (il punto fondamentale è che $\forall \mathbf{a} \in \Omega, \bigcap_{n \geq 0} K_{\mathbf{a}|n} = \{x\}$).

Consideriamo il seguente diagramma:

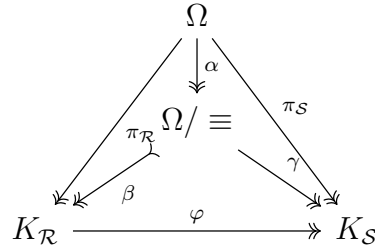


Le due proiezioni sono suriettive per il teorema fondamentale (vale anche se l'IFS non è contrattivo, a patto di avere l'ipotesi fatta sopra). Vogliamo investigare le proprietà di φ . L'assunzione cruciale è che le fibre di $\pi_{\mathcal{R}}$ siano più fini di quelle di $\pi_{\mathcal{S}}$, ossia che

$$\forall \mathbf{a}, \mathbf{a}' \in \Omega, \pi_{\mathcal{R}}(\mathbf{a}) = \pi_{\mathcal{R}}(\mathbf{a}') \implies \pi_{\mathcal{S}}(\mathbf{a}) = \pi_{\mathcal{S}}(\mathbf{a}'). \quad (*)$$

Lemma 50. *Se vale (*), allora φ è ben definita e continua. Se inoltre $\pi_{\mathcal{R}}$ e $\pi_{\mathcal{S}}$ hanno le stesse fibre, allora φ è un omeomorfismo.*

Dimostrazione. Consideriamo il seguente diagramma:



Abbiamo posto \equiv la relazione di equivalenza indotta dalle fibre di $\pi_{\mathcal{R}}$:

$$\mathbf{a} \equiv \mathbf{a}' \iff \pi_{\mathcal{R}}(\mathbf{a}) = \pi_{\mathcal{R}}(\mathbf{a}').$$

Allora è evidente che da (*) segue che γ è suriettiva.

Poniamo ora su Ω / \equiv la topologia finale. (Ricordiamo cos'è: data $(X, \tau) \xrightarrow{F} Y$ con F suriettiva, $A \subseteq Y$ è aperto nella topologia finale se e solo se $F^{-1}[A]$ è aperto in (X, τ) .) Abbiamo che Ω / \equiv è compatto in quanto immagine suriettiva di un compatto, e inoltre che β è continua. Infatti, dato $B \subseteq K_{\mathcal{R}}$ aperto, abbiamo che vale $\pi_{\mathcal{R}}^{-1}B = \alpha^{-1}(\beta^{-1}B)$ e $\beta^{-1}B$ è un aperto in Ω / \equiv per definizione di topologia finale perché $\pi_{\mathcal{R}}^{-1}B$ lo è in Ω per continuità di $\pi_{\mathcal{R}}$. Dunque β è un omeomorfismo (per un teorema visto quando abbiamo parlato di funzioni proprie) e dire che φ è continua equivale a dire che lo è γ . Questo fatto si dimostra come appena visto per β , per cui abbiamo finito.

La seconda parte del teorema è ovvia. □

Esempio 11 (scala del diavolo).

4.1 Frazioni continue e mappa di Farey

Consideriamo come al solito $\mathcal{R} = \{R_0, \dots, R_{N-1}\}$ per cui valga il teorema fondamentale degli IFS, e supponiamo di avere la seguente situazione.

$$\begin{array}{ccc} \Omega & \xrightarrow{s} & \Omega \\ \downarrow \pi & & \downarrow \pi \\ K & \xrightarrow{F} & K \end{array}$$

dove s è l'operatore di shift (ossia tronca il primo simbolo della parola infinita). Allora spesso si può definire F (almeno su qualche sottoinsieme di K) in modo che il diagramma commuti. Se ciò è possibile, allora l'unico modo è quello tale che

$$F \upharpoonright K_a = R_a^{-1}.$$

Infatti, se consideriamo $x = \pi(ab) \stackrel{TF}{=} R_a(\pi(b))$, abbiamo che

$$F(x) = F(\pi(ab)) = \pi(s(ab)) = \pi(b) = R_a^{-1}(x).$$

I problemi sorgono quando questa F non è ben definita perché per due immagini vale ad esempio $K_0 \cap K_1 \neq \emptyset$, come capita con l'IFS dovuto a Farey,...

5 Gruppo di Möbius e geometria inversiva

Sia $n \geq 1$, definiamo

$$\Pi^n := \mathbb{R}^n \cup \{\infty\} \simeq S^n$$

dove l'isomorfismo è naturalmente dato dalla proiezione stereografica.

Diamo alcune definizioni per $n = 2$, ma si possono facilmente estendere a dimensioni arbitrarie. L'*inversione di Möbius* rispetto alla retta $\langle y, w \rangle + \text{cost.} = 0$ è data da

$$x \mapsto x' = x - 2 \frac{\langle x, w \rangle}{\langle w, w \rangle} w;$$

rispetto al cerchio di centro w e raggio r è definita implicitamente da

$$\|x' - w\| \|x - w\| = r^2,$$

ed esplicitamente da

$$x \mapsto x' = w + \frac{r^2}{\|x - w\|^2} (x - w);$$

più in generale, per l'inversione rispetto a S^{n-1} in S^n si prende il cono tangente a S^n in S^{n-1} di vertice w e, preso x , si traccia la retta per x e w . L'immagine è l'altro punto di intersezione x' di tale retta con S^n .

Teorema 51. *Sia Γ la 2-sfera definita da $\|x - v\| = r$ in Π^3 . Allora l'inversione M rispetto a Γ :*

1. fissa globalmente i 2-piani per v ed agisce su di essi tramite l'inversione (nel piano) nella 1-sfera di centro v e raggio r ;
2. scambia le 2-sfere per v con i 2-piani non per v , e ciò avviene tramite proiezione stereografica rispetto a v ;
3. scambia le 2-sfere non per v con altre 2-sfere non per v ;
4. per $0 \leq m \leq 2$, conserva la classe di m -sfere e m -piani;
5. è olomorfa (ossia conserva tutti gli angoli non orientati);
6. distorce le distanze euclidee, ma conserva i birapporti;
- 1'. fissa globalmente le 2-sfere perpendicolari a Γ e agisce su di esse tramite inversione nella 1-sfera di intersezione con Γ .

Dimostrazione. Il punto (1) è ovvio. D'ora in poi senza perdita di generalità possiamo supporre che $v = 0$, ossia che $\Gamma = \{y : \|y\| = r\}$. L'espressione esplicita dell'inversione diventa quindi $M(x) = \frac{r^2}{\|x\|^2}x$. Ora, l'equazione

$$c\|x\|^2 + \langle x, w \rangle + d = 0$$

descrive sia i 2-piani che le 2-sfere. Più precisamente si tratta di un 2-piano se e solo se $c = 0$; inoltre passa per $v = 0$ se e solo se $d = 0$. Ora, se l'insieme S è definito da tale equazione, vale

$$\begin{aligned} M[S] &= \{c\|M^{-1}x\|^2 + \langle M^{-1}x, w \rangle + d = 0\} \\ &= \{c\|Mx\|^2 + \langle Mx, w \rangle + d = 0\} \\ &= \left\{ c \left\| \frac{r^2 x}{\|x\|^2} \right\|^2 + \left\langle \frac{r^2 x}{\|x\|^2}, w \right\rangle + d = 0 \right\} \\ &= \left\{ c \frac{r^4}{\|x\|^2} + \frac{r^2}{\|x\|^2} \langle x, w \rangle + d = 0 \right\} \\ &= \left\{ \frac{d}{r^2} \|x\|^2 + \langle x, w \rangle + cr^2 = 0 \right\} \end{aligned}$$

per cui la prima parte di (2) segue da questa espressione. Per quanto riguarda l'azione come proiezione stereografica, bisogna osservare che qualunque retta per v che taglia la sfera taglia anche il piano non per v in un unico punto.

Il punto (3) segue dall'espressione per $M[S]$.

Per quanto riguarda (1'), consideriamo una 2-sfera Σ perpendicolare a Γ . Se dimostriamo che Σ è fissata globalmente da M , l'affermazione segue; infatti la seconda parte è ovvia per definizione di inversione di una 2-sfera rispetto a un cerchio. Supponiamo quindi che Σ abbia centro w e raggio m . La perpendicolarità $\Sigma \perp \Gamma$ impone, per Pitagora, che

$$\|w\|^2 = r^2 + m^2.$$

Dunque Σ ha equazione $\|x - w\|^2 = \|w\|^2 - r^2$, che si traduce in

$$-\frac{1}{2}\|x\|^2 + \langle x, w \rangle - \frac{r^2}{2} = 0.$$

Quindi dall'espressione trovata prima segue che

$$\begin{aligned} M[\Sigma] &= \left\{ \frac{-r^2/2}{2}\|x\|^2 + \langle x, w \rangle + \left(-\frac{r^2}{2}\right) = 0 \right\} \\ &= \left\{ -\frac{1}{2}\|x\|^2 + \langle x, w \rangle - \frac{r^2}{2} = 0 \right\} = \Sigma \end{aligned}$$

Il punto (4) è ovvio. Per dimostrare (5) consideriamo $u \neq 0, \infty$ e due vettori nello spazio tangente $v, w \in T_u \mathbb{R}^3$. Allora M si fattorizza come

$$x \xrightarrow{Q} \|u\|^2 \frac{x}{\|x\|^2} \xrightarrow{D \frac{r^2}{\|u\|^2}} M(x),$$

quindi basta verificare che Q conservi gli angoli dato che ovviamente questo vale per la dilatazione. Si può facilmente calcolare la jacobiana di Q , che in ogni punto è simmetrica. Inoltre valgono $Q(u) = u$ e $Q^2 = \text{Id}$, per cui a livello di jacobiane vale

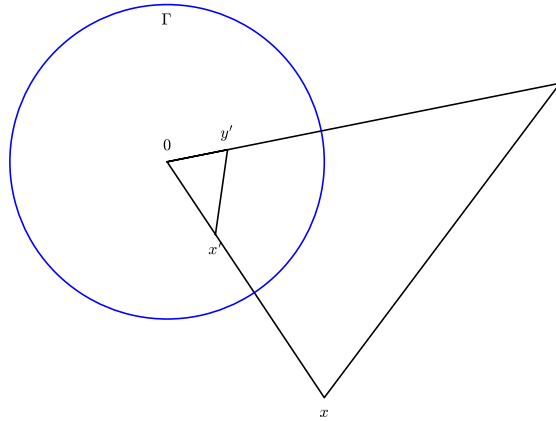
$$\text{Id} = J_{Q^2, u} = J_{Q, Q(u)} \cdot J_{Q, u} = J_{Q, u}^2,$$

quindi usando la simmetria della jacobiana

$$\langle J_{Q, u} v, J_{Q, u} w \rangle = \langle J_{Q, u}^2 v, w \rangle = \langle \text{Id} v, w \rangle = \langle v, w \rangle$$

cioè Q conserva gli angoli come volevamo.

Infine vediamo (6) considerando la costruzione in figura. I due triangoli sono simili,



avendo un angolo in comune e valendo

$$\frac{\|x\|}{\|y'\|} = \frac{\|x\|}{r^2/\|y\|} = \frac{\|y\|}{r^2/\|x\|} = \frac{\|y\|}{\|x'\|}.$$

Dunque

$$\|x' - y'\| = \|x - y\| \frac{\|y'\|}{\|x\|} = \frac{r^2}{\|x\|\|y\|} \|x - y\|,$$

per cui *in generale* la distanza euclidea non si conserva. A questo punto, scegliendo un birapporto (vale per tutti) si trova che

$$\frac{\|z' - y'\|\|x' - w'\|}{\|z' - x'\|\|y' - w'\|} = \frac{\|z - y\| \frac{r^2}{\|z\|\|y\|} \|x - w\| \frac{r^2}{\|x\|\|w\|}}{\|z - x\| \frac{r^2}{\|z\|\|x\|} \|y - w\| \frac{r^2}{\|y\|\|w\|}} = \frac{\|z - y\|\|x - w\|}{\|z - x\|\|y - w\|}.$$

□

Ci sono due conseguenze del risultato appena ottenuto. La prima è che dato un *cerchio esteso* (ossia un vero cerchio o una retta) C e un punto $x \notin C$, l'immagine dell'inversione rispetto a tale cerchio può essere calcolata come l'intersezione della retta che congiunge x con il centro di C con l'unico cerchio (esteso) passante per x e ortogonale a C . Infatti entrambi i cerchi estesi per x sono fissati globalmente e si intersecano solo in x e x' .

La seconda conseguenza è che la proiezione stereografica coniuga le due definizioni di inversione, quella su Σ^n e quella su Π^n .

(Fare disegni).

Definizione 52. Dato $n \geq 1$ e fissato di conseguenza S^n (o equivalentemente Π^n), il *gruppo di Möbius* Möb_n è il gruppo di biezioni di S^n generato dalle inversioni nelle n -sfere perpendicolari al bordo di S^n (o nel caso di Π^n , nelle $(n - 1)$ -sfere in cui queste intersecano S^n).

Osservazione 53. Il gruppo di Möbius comprende traslazioni, rotazioni, simmetrie assiali e dilatazioni. Pensando in Π^n è facile capirlo dato che l'inversione rispetto a rette non è altro che la simmetria assiale e con questa si generano rotazioni e traslazioni; le dilatazioni si ottengono invertendo rispetto a due cerchi concentrici. Ha un sottogruppo di indice 2, Möb_n^+ , che contiene gli elementi che preservano l'orientazione e si ottiene in modo ovvio moltiplicando un numero pari di elementi di Möb_n .

Concentriamoci ora sul caso di $n = 2$. Dati $z \in \mathbb{C} \cup \{\infty\}$ e

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in PSL_2\mathbb{C},$$

l'*omografia indotta da A* è la mappa

$$(A, 0)(z) = \frac{\alpha z + \beta}{\gamma z + \delta}$$

mentre l'*antiomografia indotta da A* è

$$(A, 1)(z) = \frac{\alpha \bar{z} + \beta}{\gamma \bar{z} + \delta}.$$

Risulta evidente che valga il seguente schema per le composizioni di due [anti]omografie:

$$\begin{aligned} (A, 0) \circ (B, 0) &= (AB, 0) \\ (A, 0) \circ (B, 1) &= (AB, 1) \\ (A, 1) \circ (B, 0) &= (A\bar{B}, 1) \\ (A, 1) \circ (B, 1) &= (A\bar{B}, 0) \end{aligned}$$

per cui il gruppo di omografie e antiomografie è $PSL_2\mathbb{C} \rtimes \mathbb{Z}_2$.

Osservazione 54. Fissato un cap \mathcal{H}^n di S^n , sia $\text{stab}(\mathcal{H}^n, \text{Möb}_n)$ il suo stabilizzatore in Möb_n . Una volta dimostrato che tale stabilizzatore è generato dalle inversioni rispetto a cerchi perpendicolari al bordo di \mathcal{H}^n , dato che queste generano Möb_{n-1} risulterà $\text{stab}(\mathcal{H}^n, \text{Möb}_n) \simeq \text{Möb}_{n-1}$.

Lemma 52 (decomposizione di Iwasawa). *Ogni $M \in SL_2\mathbb{C}$ è scrivibile in modo unico come $M = KAN$ dove $K \in \mathcal{K} = SU_2\mathbb{C}$, $A \in \mathcal{U} = \left\{ \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}, a > 0 \right\}$ e $N \in \mathcal{N} = \left\{ \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}, \beta \in \mathbb{C} \right\}$.*

Dimostrazione. Consideriamo l'immagine della base canonica tramite M ,

$$(e_1, e_2)M = (m_1, m_2)$$

e fissato $u_1 = m_1$ imponiamo l'ortogonalità con $u_2 = m_2 - \alpha m_1$, ossia $\langle m_1, m_2 - \alpha m_1 \rangle = 0$, che determina $\alpha = \langle m_1, m_2 \rangle / \langle m_1, m_1 \rangle$. Abbiamo quindi la seguente relazione

$$(u_1, u_2) = (m_1, m_2) \underbrace{\begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}}_{=: N^{-1}}.$$

La nuova base è ortogonale, quindi poiché

$$(e_1, e_2)MN^{-1} = (u_1, u_2)$$

la matrice MN^{-1} è ortogonale (non per forza hermitiana!). Di conseguenza esiste $a > 0$ tale che

$$(MN^{-1})^*(MN^{-1}) = \begin{pmatrix} a^2 & 0 \\ 0 & a^{-2} \end{pmatrix} =: A^2$$

per cui

$$(MN^{-1}A^{-1})^*(MN^{-1}A^{-1}) = I_2$$

e quindi $MN^{-1}A^{-1} =: K \in \mathcal{K}$, e abbiamo dimostrato l'esistenza.

Per quanto riguarda l'unicità della decomposizione, supponiamo che $K_1A_1N_1 = K_2A_2N_2$. Allora

$$N_1^*A_1^*\underbrace{K_1^*K_1}_{=I_2}A_1N_1 = N_2^*A_2^*\underbrace{K_2^*K_2}_{=I_2}A_2N_2$$

per cui dal fatto che le A_i sono diagonali a elementi reali si ha $A_i^* = A_i$, quindi

$$N_1^*A_1^2N_1 = N_2^*A_2^2N_2.$$

A questo punto moltiplichiamo a sinistra per $(N_2^*)^{-1}$ e a destra per N_1^{-1} , ottenendo

$$(N_2^*)^{-1}N_1^*A_1^2 = A_2^2N_2N_1^{-1}$$

e qui il termine a sinistra è triangolare inferiore, quello a destra triangolare superiore. Di conseguenza entrambi devono in realtà essere diagonali, da cui seguono $N_1 = N_2$ e $A_1 = A_2$. \square

Teorema 53. Vale $\text{Möb}_2 = \text{PSL}_2\mathbb{C} \rtimes \mathbb{Z}_2$, e di conseguenza $\text{Möb}_2^+ = \text{PSL}_2\mathbb{C}$.

Dimostrazione. (\subseteq) L'inversione rispetto alla circonferenza unitaria centrata nell'origine è data da

$$z \mapsto \frac{1}{\bar{z}} = \left(\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, 1 \right)$$

per cui è un'antimografia. La dilatazione di un generico fattore $r > 0$ è data da

$$\left(\begin{pmatrix} r^{1/2} & 0 \\ 0 & r^{-1/2} \end{pmatrix}, 0 \right),$$

la traslazione di $\alpha \in \mathbb{C}$ da

$$\left(\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, 0 \right)$$

e quindi per le inversioni rispetto a cerchi va tutto bene. Per quanto riguarda le inversioni rispetto a rette, sono tutte coniugate (a meno di rototraslazione) a quella rispetto all'asse reale, e questa è indotta da

$$\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, 1 \right).$$

(\supseteq) Si utilizza la decomposizione di Iwasawa per elementi di $SL_2\mathbb{C}$, per cui preso un generico $M \in SL_2\mathbb{C}$ abbiamo $M = KAN$. Abbiamo che $A, N \in \text{Möb}_2$, per cui ci resta da determinarlo per $K \in SU_2\mathbb{C}$. Studiamo meglio la forma di $K = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. Dato che sta in $SU_2\mathbb{C}$, deve valere $K^{-1} = K^*$ che si traduce in termini di elementi come $\delta = \bar{\alpha}$ e $\gamma = -\bar{\beta}$, per cui abbiamo che

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad \det K = |\alpha|^2 + |\beta|^2 = 1.$$

Se $K \neq \pm \text{Id}$, ci sono due autovalori complessi coniugati di modulo 1 e due corrispondenti autovettori.

Osserviamo ora che $\forall p, q \in P^1\mathbb{C}$ esiste un elemento B di Möb_2 che manda $p \mapsto \infty$ e $q \mapsto 1$. Infatti se $p = \infty$ basta applicare una traslazione, se $p \neq \infty$ si considera l'inversione rispetto alla circonferenza di centro p e passante per q e si ricade nel caso precedente.

Possiamo quindi supporre senza perdita di generalità che gli autovettori di K siano 1 e ∞ . L'equazione che testimonia questo fatto è

$$BKB^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ 0 & \bar{\rho} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

e osserviamo che la penultima matrice è una rotazione di angolo $2\arg(\rho)$, che sta in Möb_2 . Quindi $K \in \text{Möb}_2$. \square

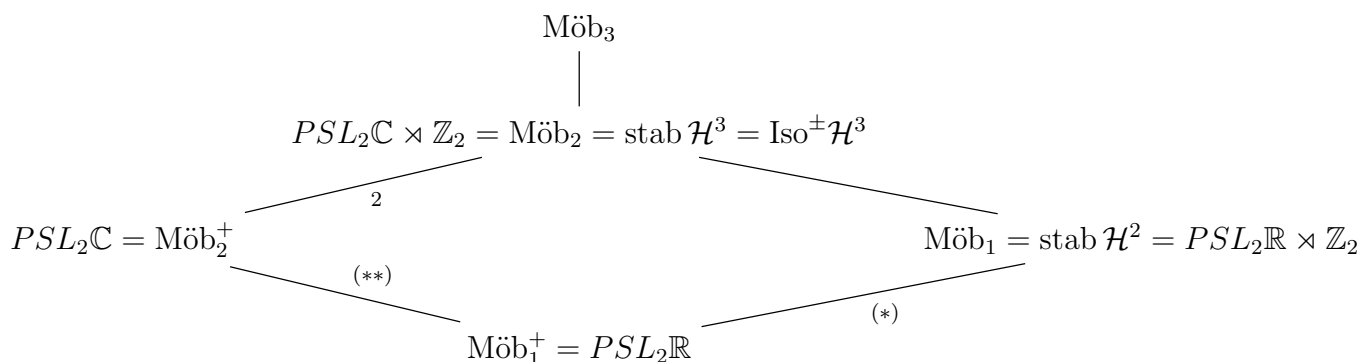
Lo spazio iperbolico \mathcal{H}^3 si può pensare in due modi: come l'interno di una sfera S^2 (che sarebbe un cap di S^3) oppure grazie alla proiezione stereografica come la parte "sopra" al piano $\mathbb{C} \cup \{\infty\}$. La distanza che si usa è quella iperbolica, definita inizialmente per il modello di Klein di \mathcal{H}^2 ma che va bene anche per gli altri casi, data da

$$d_B(p_1, p_2) = \log \frac{\|p_0 - p_2\| \|p_1 - p_3\|}{\|p_0 - p_1\| \|p_2 - p_3\|}$$

dove p_0 e p_3 sono i punti che si ottengono facendo l'intersezione dell'unico 1-cerchio per p_1 e p_2 perpendicolare a $\partial\mathcal{H}^2$ (o $\partial\mathcal{H}^3$) con tale bordo. Poiché le inversioni conservano il birapporto, gli elementi di Möb_2 (o Möb_3) sono isometrie per lo spazio iperbolico. Precisiamo questo fatto nel seguente teorema.

Teorema 54. *Ogni isometria di \mathcal{H}^3 è nel sottogruppo di Möb_3 generato dalle inversioni nelle 2-sfere perpendicolari a $\partial\mathcal{H}^3$ (più in generale per il caso di \mathcal{H}^n bastano $n+1$ inversioni). In particolare, tale sottogruppo è proprio $\text{stab}(\mathcal{H}^3, \text{Möb}_3)$.*

Schema di dimostrazione. Per omogeneità, un'isometria è determinata dalla sua azione su un qualsiasi simpleso dello spazio. Preso dunque un tetraedro di vertici p_0, p_1, p_2, p_3 e un'isometria Φ , diciamo che le immagini dei vertici del tetraedro siano q_0, q_1, q_2, q_3 . Si manda dunque $p_0 \mapsto q_0$ con un'opportuna inversione (questo è facile!). Dopodiché si prende un'inversione che manda $p_1 \mapsto q_1$ ma la cui sfera contiene q_0 , che quindi resta fissato, e così avanti. Si riesce giusto in tempo perché c'è solo una sfera per tre punti e perpendicolare al bordo, ma è proprio quella che scambia tra loro i due possibili tetraedri costruiti sulle immagini dei primi tre punti. \square



Osservazione 55. Consideriamo un generico $z \in \mathbb{C}$. Esso appartiene a \mathcal{H}^2 se e solo se

$$\begin{pmatrix} z \\ 1 \end{pmatrix}^* \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} z \\ 1 \end{pmatrix} < 0.$$

Infatti svolgendo i conti si ottiene che tale condizione equivale a

$$-i\bar{z} + iz = i(z - \bar{z}) = i2\Im z \cdot i = -2\Im z < 0.$$

Ne segue che una generica matrice $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in PSL_2\mathbb{C}$ stabilizza globalmente il piano iperbolico \mathcal{H}^2 (considerato nel modello del semipiano) se e solo se lascia invariata tale forma, ossia

$$A^* \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} A = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

(In realtà questa relazione dovrebbe valere a meno di costante moltiplicativa perché quello che ci interessa è soltanto che il segno della forma resti invariato, ma la costante moltiplicativa si può sempre assorbire nella matrice di $PSL_2\mathbb{C}$). Svolgendo i calcoli si ottiene che ciò equivale a $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, che dimostra **(**)** nel reticolo qui sopra.

Per tornare al piano di sopra dall'altro lato (*), bisogna aggiungere l'inversione $z \mapsto -\bar{z}$, che è indotta da

$$\left(\begin{pmatrix} i & 0 \\ 0 & i^{-1} \end{pmatrix}, 1 \right)$$

ma anche da

$$\left(\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, 1 \right)$$

che è una scelta più furba perché permette di non uscire dai reali.

5.1 Gruppi fuchsiani/kleiniani

Lemma 55. *Sia \mathcal{G} un gruppo che agisce in modo transitivo sullo spazio topologico di Hausdorff localmente compatto X . Le due condizioni seguenti sono equivalenti e definiscono un'azione propriamente discontinua:*

1. $\forall x \in X, \forall K \subseteq X$ compatto, $\#\{G \in \mathcal{G} : Gx \in K\} < \infty$;
2. l'orbita di ogni $x \in X$ è discreta e ogni punto ha stabilizzatore finito.

Dimostrazione. (1 \implies 2) Che lo stabilizzatore sia finito si ottiene facilmente scegliendo $K = \{x\}$. Ora, visto che l'azione è transitiva, basta dimostrare che $\forall x \in X, \exists O \ni x$ aperto con $O \cap \mathcal{G} * x = \{x\}$. Per la locale compattezza, questo O si può scegliere in modo che la sua chiusura sia compatta. Per (1) però questo implica che tale chiusura contenga un numero finito di punti dell'orbita. Quindi qui fissato un particolare punto dell'orbita, esso si può separare dagli altri grazie all'ipotesi che lo spazio sia di Hausdorff.

(2 \implies 1) Dati x e K , per ipotesi abbiamo che $\mathcal{G} * x$ è discreta. Poiché K è compatto, l'intersezione $\mathcal{G} * x \cap K$ è finita. Infine la cardinalità dell'insieme delle mappe da x a un certo y è la stessa dello stabilizzatore di y , che per ipotesi è finito. Di conseguenza è finita anche la cardinalità dell'insieme delle mappe che mandano x nei punti di $\mathcal{G} * x \cap K$, in quanto unione finita di insiemi finiti. \square

Osservazione 56. Se \mathcal{G} è anche un gruppo topologico e l'azione è continua e propria (cioè i controinsiemi di compatti sono compatti, rispetto alla mappa $G \times X \rightarrow X$) e consideriamo un sottogruppo $\Gamma \leq \mathcal{G}$ discreto, allora l'azione di Γ su X è propriamente discontinua.

Dimostrazione. Verifichiamo la condizione 1 del lemma, ossia che dati x e K , l'insieme $\{G \in \Gamma : Gx \in K\}$ è finito. Tale insieme si può riscrivere come

$$\Gamma \cap \{G \in \mathcal{G} : Gx \in K\} = \Gamma \cap \{(G, y) \in \mathcal{G} \times X : Gy \in K\} \cap \pi^{-1}\{x\}$$

e siccome l'azione è propria, $\{(G, y) : Gy \in K\}$ è compatto; inoltre $\pi^{-1}\{x\}$ è chiuso, per cui anche l'intersezione $\{(G, y) : Gy \in K\} \cap \pi^{-1}\{x\}$ è compatta essendo un chiuso in un compatto. Infine, essendo Γ discreto, ovviamente intersecandolo con un compatto si ottiene un insieme finito. \square

Definizione 57. Un sottogruppo Γ di Möb_{n-1} che agisce in modo propriamente discontinuo su \mathcal{H}^n si dice un *gruppo fuchsiano esteso* se $n = 2$ (ossia è un sottogruppo di $PSL_2\mathbb{R} \rtimes \mathbb{Z}_2$), un *gruppo kleiniano esteso* se $n = 3$ (sottogruppo di $PSL_2\mathbb{C} \rtimes \mathbb{Z}_2$). I loro sottogruppi che preservano l'orientazione si dicono *gruppi fuchsiani* e *kleiniani* rispettivamente.

Esempio 12. $PSL_2\mathbb{Z}$ e $PSL_2\mathbb{Z} \rtimes \mathbb{Z}_2$ sono fuchsiani; il gruppo modulare di Picard $PSL_2\mathbb{Z}[i]$ è kleiniano. Esempi più generali si hanno considerando i vari $PSL_2\mathbb{O}$, dove \mathbb{O} è l'anello degli interi algebrici di una fissata estensione quadratica complessa di \mathbb{Q} , detti gruppi di Bianchi. Le estensioni quadratiche *reali* non vanno bene, perché se ad esempio abbiamo $\mathbb{Q}(\lambda)$ di grado 2, vale

$$\begin{pmatrix} 1 & \lambda^{-k} \\ 0 & 1 \end{pmatrix} \xrightarrow{k \rightarrow \infty} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

quindi l'identità I_2 è punto di accumulazione, per cui non si possono avere orbite discrete.

5.2 Classificazione inversioni

Abbiamo visto che i gruppi Mob_1^+ e Mob_2^+ sono isomorfi a $PSL_2\mathbb{R}$ e $PSL_2\mathbb{C}$ rispettivamente. Vogliamo classificare le trasformazioni di tali gruppi, e lo facciamo per il più grande (la classificazione delle trasformazioni di $PSL_2\mathbb{R}$ sarà una conseguenza di questa).

Consideriamo quindi una generica matrice

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in PSL_2\mathbb{C}.$$

Essa agisce in modo naturale su $P^1\mathbb{C}$ e per estensione anche su \mathcal{H}^3 pensato come lo spazio "sopra" a $P^1\mathbb{C}$.

Osservazione 58. In effetti possiamo pensare ad \mathcal{H}^3 come ad un sottoinsieme dei quaternioni \mathbb{H} , in particolare quello dei quaternioni della forma

$$q = z + rj, \quad z \in \mathbb{C}, r > 0.$$

A questo punto l'inversione rispetto ad una circonferenza di $P^1\mathbb{C}$ agisce su \mathcal{H}^3 come l'inversione rispetto alla (unica) sfera che ha tale circonferenza come equatore. Algebricamente, ricordando che i quaternioni non sono commutativi, l'azione della matrice A definita sopra si esprime come

$$A * q = (\alpha q + \beta)(\gamma q + \delta)^{-1}.$$

Vediamo che la classificazione di tali trasformazioni dipende essenzialmente dalla traccia della matrice, $t := \alpha + \delta$. Essendo il polinomio caratteristico dato da $x^2 - tx + 1$, gli autovalori di A sono $\frac{t \pm \sqrt{t^2 - 4}}{2}$. Distinguiamo alcuni casi.

- $t^2 = 4$, ovvero $t = \pm 2$. Abbiamo allora un unico autovalore che vale ± 1 e A risulta coniugata a

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

che (sul piano) agisce come una traslazione euclidea. A in questo caso si dice *parabolica*.

- $t \in (-2, 2) \subset \mathbb{R}$. Allora gli autovalori sono ρ e $\bar{\rho}$ e la matrice A è coniugata a

$$\begin{pmatrix} \rho & 0 \\ 0 & \bar{\rho} \end{pmatrix}$$

che agisce sul piano come una rotazione $z \mapsto \rho^2 z$. A si dice *ellittica*.

- $t \in \mathbb{C} \setminus [-2, 2]$. Allora gli autovalori sono ρ e ρ^{-1} con $|\rho| > 1$, A è coniugata a

$$\begin{pmatrix} \rho & 0 \\ 0 & \rho^{-1} \end{pmatrix}$$

e agisce come una rotodilatazione $z \mapsto \rho^2 z$. A si dice *lossodromica*. Un sottocaso particolare di questo si ha quando $t \in \mathbb{R} \setminus [-2, 2]$, nel qual caso A si dice *iperbolica* (e infatti per le matrici di $PSL_2\mathbb{R}$ di solito si parla solo di queste ultime).

5.3 Insiemi limite

Definizione 59. Sia $\Gamma < PSL_2\mathbb{C}$ kleiniano. Il suo *insieme limite* è

$$\Lambda = \Lambda(\Gamma) = \{p \in \partial\mathcal{H}^3 : \exists q \in \mathcal{H}^3 \text{ t.c. } p \text{ è di accumulazione per l'orbita } \Gamma * q \text{ di } q\}.$$

Importante: nella definizione appena data, l'accumulazione si intende rispetto alla norma euclidea.

Osservazione 60. Se p è di accumulazione per $\Gamma * q$, allora lo è pure per $\Gamma * t$.

Dimostrazione. Se p è di accumulazione per $\Gamma * q$, allora esiste una successione $G_0, G_1, \dots \in \Gamma$ tale che $G_n * q \rightarrow p$ in senso euclideo per $n \rightarrow \infty$. Consideriamo l'orbita $G_n * t$. Poiché le G_n stanno in $PSL_2\mathbb{C}$, esse conservano la distanza iperbolica. Quindi se d è la distanza iperbolica tra q e t , è anche la distanza iperbolica tra $G_n q$ e $G_n t$ per ogni n . Però siccome il punto di accumulazione p di $\Gamma * q$ sta sul bordo $\partial\mathcal{H}^3$, tali distanze in senso euclideo vanno a 0, per cui si può concludere per la disuguaglianza triangolare che $G_n * t \rightarrow p$. \square

Osservazione 61. Ovviamente Λ è un chiuso perché il suo complementare $\Omega = \partial\mathcal{H}^3 \setminus \Lambda$, detto anche *insieme regolare* o *dominio di discontinuità* di Γ è aperto. Altrettanto ovviamente Λ è Γ -invariante, e gli elementi di Γ agiscono su Λ come omeomorfismi. Una proprietà meno ovvia è che se Λ non ricade in alcuni casi speciali (vuoto, un punto, due punti, $\partial\mathcal{H}^3$) allora è un *perfetto* (chiuso senza punti di accumulazione) ad interno vuoto.

Esempio 13. Le rotazioni non hanno punti di accumulazione sul bordo; le lossodromiche ne hanno due; il gruppo di inversioni generato da quelle in cerchi con diametri tutti lungo una circonferenza massima di $\partial\mathcal{H}^3$ ha come insieme limite la circonferenza massima stessa; il gruppo di Picard ha (probabilmente) $\Lambda = \partial\mathcal{H}^3$.

5.4 Unificazione di cap e punti

Consideriamo ora il gruppo (fuchsiano esteso) generato dalle inversioni rispetto a tre cerchi disgiunti in $P_1\mathbb{C}$ (o in alternativa in S^2). Vogliamo studiare il suo comportamento utilizzando il chaos game. Siccome le trasformazioni che andiamo ad applicare sono involuzioni, è importante non applicare due volte consecutive la stessa. Questo obiettivo si può raggiungere facilmente se l'indice della "nuova" trasformazione da applicare si ottiene sommando 1 o 2 e poi facendo mod 3.

Se usassimo un gruppo fuchsiano tipo

$$\Gamma = \langle A_0, \dots, A_{N-1}, A_0^{-1}, \dots, A_{N-1}^{-1} \rangle$$

avremmo il problema simile di evitare di applicare una trasformazione e la sua inversa consecutivamente.

Vediamo ora una interessante costruzione che permette di rendere tutto più semplice. Consideriamo la sfera S^n . Ad un suo cap C si può associare il suo centro $c \in S^n$ e l'angolo di apertura $\theta \in (0, \pi)$. Allora il cap si può caratterizzare come

$$\begin{aligned} C &= \{x \in S^n : \langle x, c \rangle \geq \cos \theta\} = \{x \in S^n : \langle x, c \rangle - \cos \theta \geq 0\} \\ &= \{x \in S^n : \langle x, \frac{c}{\sin \theta} \rangle - \cot \theta \geq 0\}. \end{aligned}$$

A questo punto descriviamo sia i punti che i caps come dei vettori in \mathbb{R}^{n+2} tramite le seguenti associazioni:

- $x \in S^n$ viene associato a $\mathbf{x} = (x_1, \dots, x_{n+1}, 1)^t$;
- $C \sim (c, \theta)$ viene associato a $\mathbf{w} = (c_1/\sin \theta, \dots, c_{n+1}/\sin \theta, \cot \theta)^t$.

D'ora in avanti useremo $\langle -, - \rangle$ anche per vettori in \mathbb{R}^{n+2} indicando il prodotto scalare di Lorentz di segnatura $(n+1, 1)$. Dal contesto (vettori in grassetto) sarà sempre chiaro quale stiamo usando.

Con queste convenzioni, abbiamo le seguenti conseguenze:

- $x \in C$ se e solo se $\langle \mathbf{x}, \mathbf{w} \rangle \geq 0$;
- $\mathbf{x} \in \mathbb{R}^{n+2}$ corrisponde ad un punto se e solo se $\mathbf{x}_{n+1} = 1$ e $\langle \mathbf{x}, \mathbf{x} \rangle = 0$;
- $\mathbf{w} \in \mathbb{R}^{n+2}$ corrisponde ad un cap se e solo se appartiene allo *spazio di De Sitter* $S = \{\mathbf{z} \in \mathbb{R}^{n+2} : \langle \mathbf{z}, \mathbf{z} \rangle = 1\}$.

Dimostrazione. Dimostriamo solo l'ultima equivalenza. (\Rightarrow) è immediato:

$$\left(\frac{c_1}{\sin \theta}\right)^2 + \dots + \left(\frac{c_{n+1}}{\sin \theta}\right)^2 - (\cot \theta)^2 = \frac{1 - (\cos \theta)^2}{(\sin \theta)^2} = 1.$$

Vediamo (\Leftarrow) . Sia \mathbf{y} tale che $\langle \mathbf{y}, \mathbf{y} \rangle = 1$. Basta porre $\theta = \operatorname{arccot} y_{n+2}$ e di conseguenza, dato che il centro del cap deve appartenere alla sfera, $c := (y_1, \dots, y_{n+1})/(1+y_{n+2}^2)^{1/2}$. \square

La magia di tutta questa costruzione sta nel seguente risultato.

Teorema 56. *L'inversione rispetto a $C \sim \mathbf{w}$ è data da*

$$\mathbf{x} \mapsto \mathbf{x} - 2\langle \mathbf{w}, \mathbf{x} \rangle \cdot \mathbf{w}$$

Tale mappa se consideriamo come ambiente S^2 è un elemento di $GL_4\mathbb{R}$, più precisamente di $O_{3,1}\mathbb{R}$. La cosa importante è che possiamo scriverla in modo esplicito(!), infatti preso un punto $\mathbf{x} = (x_1, x_2, x_3, 1)^t$ la sua immagine è

$$I\mathbf{x} - 2\mathbf{w}\mathbf{w}^t \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} \mathbf{x} = [I - 2\mathbf{w}\mathbf{w}^t L]\mathbf{x}.$$

Esercizio in SageMath: consideriamo inizialmente i tre cap corrispondenti a $\mathbf{w}_0 = (0, 1, -1, 1)^t$, $\mathbf{w}_1 = (0, -1, -1, 1)^t$, $\mathbf{w}_2 = (0, -1, 1, 1)^t$ e usiamo il chaos game su questi. Infine aggiungiamo un cap parametrico $\mathbf{w}_4 = (1/\sin \theta, 0, 0, \cot \theta)$ e facciamo variare θ facendolo crescere fino a quando il cap corrispondente non è tangente agli altri tre.

Teorema 57. *L'inversione nel cap $C \sim \mathbf{w}$ è data dalla mappa*

$$\mathbf{x} \mapsto \mathbf{x} - 2 \frac{\langle \mathbf{w}, \mathbf{x} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}.$$

Tale mappa è un elemento di $O_{3,1}\mathbb{R}$, e tale espressione funziona sia per i punti che per i cap.

Dimostrazione. Dati S^3 e \mathbf{w} , osserviamo che $\langle \mathbf{x}, \mathbf{w} \rangle = 0$ individua il piano polare a \mathbf{w} rispetto alla conica S^3 . Allora preso un generico \mathbf{x} , possiamo scriverlo in modo unico come $\mathbf{x} = \mathbf{u} + r\mathbf{w}$ con \mathbf{u} in tale piano polare. La sua immagine tramite la mappa definita nell'enunciato è $\mathbf{x}' = \mathbf{u} - r\mathbf{w}$. Osserviamo che $\mathbf{x} \mapsto \mathbf{x}'$ conserva la forma di Lorentz (esercizio), per cui se dimostriamo che i tre punti x, x', w (corrispondenti nello spazio di partenza di \mathbf{x}, \mathbf{x}' , con l'ultimo che è il "vertice" del cap corrispondente a \mathbf{w}) sono allineati allora x' è proprio l'immagine di x tramite l'inversione nel cap $C_{\mathbf{w}}$, in quanto valgono $\langle \mathbf{x}', \mathbf{x}' \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = 0$ e l'intersezione di una retta con S^2 ha due punti (ricordare la costruzione geometrica dell'inversione!). Che siano allineati si ottiene considerando il fatto che $\mathbf{x}' = \mathbf{x} - 2 \frac{\langle \mathbf{w}, \mathbf{x} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}$, per cui $\mathbf{x}', \mathbf{x}, \mathbf{w}$ stanno sullo stesso 2-piano e quindi x, x', w sulla stessa retta proiettiva.

La stessa formula funziona anche per i cap perché conserva la forma di Lorentz e quindi manda cap in cap. Inoltre, per lo stesso motivo, manda $\mathbf{x} \in C_{\mathbf{w}}$ in $\mathbf{x}' \in C_{\mathbf{w}'}$, dato che l'appartenenza ad un cap si esprime in termini di forma di Lorentz. \square

Lemma 58. *Sia $A \in O_{3,1}\mathbb{R}$. Allora $a_{44} \neq 0$ e A fissa globalmente ognuna delle due componenti connesse $\pm\mathcal{L}$ di $\{\mathbf{x} : \langle \mathbf{x}, \mathbf{x} \rangle = -1\}$ oppure le scambia a seconda del segno di a_{44} . Di conseguenza l'omomorfismo*

$$O_{3,1}\mathbb{R} \rightarrow \mathbb{Z}_2 \times \mathbb{Z}_2, \quad A \mapsto (\det A, \text{sgn} a_{44})$$

induce l'affettamento di $O_{3,1}\mathbb{R}$ in

$$\begin{array}{ccc} \text{sgn} = 1 & & \\ \det = 1 & SO_{3,1}^\uparrow \mathbb{R} & \begin{pmatrix} 1 & & & -1 \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} SO_{3,1}^\uparrow \mathbb{R} \\ -1 & \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} SO_{3,1}^\uparrow \mathbb{R} & \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} SO_{3,1}^\uparrow \mathbb{R} \end{array}$$

con la prima riga che forma tutto $SO_{3,1}\mathbb{R}$ e la prima colonna che forma $O_{3,1}^\uparrow \mathbb{R}$. La freccia in su indica che il gruppo non scambia $\pm\mathcal{L}$.

Dimostrazione. Siano $\mathbf{v} = (0, 0, 0, 1)^t \in \mathcal{L}$ e $A \in O_{3,1}$. Per definizione la forma di Lorentz viene conservata da A , per cui vale

$$-1 = \langle \mathbf{v}, \mathbf{v} \rangle = \langle A\mathbf{v}, A\mathbf{v} \rangle = a_{14}^2 + a_{24}^2 + a_{34}^2 - a_{44}^2$$

e quindi necessariamente deve essere $a_{44} \neq 0$.

Infine per vedere se le due componenti $\pm \mathcal{L}$ vengono scambiate o meno basta (per connessione e continuità) vedere in quale delle due falde sta l'immagine di \mathbf{v} , ossia valutare la sua quarta coordinata a_{44} . La tesi è a questo punto ovvia. \square

Teorema 59. *La mappa che ad ogni inversione in $P^1\mathbb{C}$ rispetto al cap $C \sim \mathbf{w}$ associa la matrice $R_{\mathbf{w}} \in O_{3,1}\mathbb{R}$ che induce la trasformazione del Teorema 53, esplicitamente data da $R_{\mathbf{w}} = I - 2\mathbf{w}\mathbf{w}^t L$ induce due isomorfismi*

$$PSL_2\mathbb{C} \rightarrow SO_{3,1}^\uparrow\mathbb{R}, \quad PSL_2\mathbb{C} \rtimes \mathbb{Z}_2 \rightarrow O_{3,1}^\uparrow\mathbb{R}.$$

Dimostrazione. Per prima cosa osserviamo che per il teorema precedente la matrice $R_{\mathbf{w}}$ preserva \mathcal{L} in quanto $(R_{\mathbf{w}})_{4,4} = 1 - 2w_4^2 \cdot (-1) = 1 + 2w_4^2 > 0$.

Che la mappa descritta induca omomorfismi è ovvio. Abbiamo che l'inversione $(I, 1) : z \mapsto \bar{z}$ corrisponde a $R_{\mathbf{e}_2}$: è l'inversione rispetto al cerchio massimo di S^2 "parallelo all'asse reale". Per dimostrare che l'omomorfismo è iniettivo supponiamo per assurdo che $(A, 0/1) \mapsto I_{\mathbb{R}^4}$. Allora l'immagine in particolare fissa globalmente $\{\mathbf{x} : \langle \mathbf{x}, \mathbf{x} \rangle = 0\} = S^2$ per cui in particolare $A = \pm I$ e, dato che nessuna antiomotetia è l'identità su S^2 , deve essere $A = I$.

Infine resta da verificare la suriettività. Facciamo un giro un po' lunghetto. Consideriamo $\mathfrak{w} := \{W \in \mathbb{C}^{2 \times 2} : W^* = -W\} = \mathfrak{u}_2\mathbb{C}$. Facendo i conti si scopre che le matrici di \mathfrak{w} hanno la forma

$$W = \begin{pmatrix} ri & \beta \\ -\bar{\beta} & si \end{pmatrix}, \quad \beta \in \mathbb{C}, \quad r, s \in \mathbb{R}$$

per cui abbiamo quattro gradi di libertà (reali).

Ora consideriamo la seguente operazione: prendiamo $A \in PSL_2\mathbb{C}$ e coniughiamo W per A^{-1} . Allora si rimane in \mathfrak{w} , infatti

$$[(A^{-1})^* W A^{-1}]^* = (A^{-1})^* W^* A^{-1} = -[(A^{-1})^* W A^{-1}].$$

A questo punto, se $\mathbf{w} = (w_1, w_2, w_3, w_4)$ prendendo la seguente matrice di \mathfrak{w}

$$W = \begin{pmatrix} (w_3 - w_4)i & -w_2 + w_1i \\ w_2 + w_1i & -(w_3 + w_4)i \end{pmatrix},$$

si ha che il coniugio definito sopra è anche lineare rispetto a W .

A questo punto usiamo un fatto che non dimostriamo: nella proiezione stereografica si ha $\mathbf{x} \in C_{\mathbf{w}}$ se e solo se $\langle \mathbf{x}, \mathbf{w} \rangle > 0$ (e fin qui nulla di nuovo) se e solo se

$$\begin{pmatrix} \alpha \\ 1 \end{pmatrix}^* W \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \in \mathbb{R}_{>0} \cdot i,$$

dove α è l'immagine di \mathbf{x} rispetto alla proiezione stereografica.

Questo ci torna utile perché si può caratterizzare l'appartenenza ad un certo disco come segue: $\alpha \in \text{disco}$ se e solo se $\begin{pmatrix} \alpha \\ 1 \end{pmatrix}^* D \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \in \mathbb{R}_{>0} \cdot i$ per una opportuna matrice D . Ma inserendo $A^*(A^*)^{-1}$ e AA^{-1} , abbiamo che quest'ultima condizione è equivalente a

$$\left[A \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \right]^* (A^*)^{-1} D A^{-1} \left[A \begin{pmatrix} \alpha \\ 1 \end{pmatrix} \right] \in \mathbb{R}_{>0} \cdot i,$$

per cui $(A^*)^{-1}DA^{-1}$ caratterizza il disco di arrivo.

A questo punto si riconosce che $\mathbf{w} \mapsto \mathbf{A}\mathbf{w} = \mathbf{w}'$ se e solo se

$$(A^{-1})^*WA^{-1} = \begin{pmatrix} (w'_3 - w'_4)i & -w'_2 + w'_1i \\ w'_2 + w'_1i & -(w'_3 + w'_4)i \end{pmatrix}$$

che implicitamente definisce la matrice $A \in PSL_2\mathbb{C}$ che stavamo cercando. □