

Teoria ergodica ed entropia

Alessio Muscillo

Indice

1	Trasformazioni che preservano la misura	5
1.1	Ricorrenza	5
1.2	Convergenza in media	8
1.3	Convergenza puntuale e teoremi ergodici	10
1.4	Ergodicità	18
1.5	Mixing e altre proprietà	27
2	Isomorfismi e spettro di trasformazioni	33
2.1	Spettro continuo e proprietà ergodiche	33
2.2	Trasformazioni isomorfe e coniugate	36
2.3	Isomorfismi spettrali e invarianti spettrali	38
3	Entropia	43
3.1	Nozioni preliminari	43
3.2	Entropia	44
3.3	Entropia condizionale	46
3.4	Alcune proprietà di $h(T, \mathcal{A})$ e di $h(T)$	49
3.5	Calcolo dell'entropia	54
3.6	β -espansioni	56
	Bibliografia	59

Capitolo 1

Trasformazioni che preservano la misura

1.1 Ricorrenza

Definizione 1.1.1. Siano $(X_1, \mathcal{B}_1, m_1)$ e $(X_2, \mathcal{B}_2, m_2)$ spazi di misura, e sia $T : X_1 \rightarrow X_2$. Diciamo che:

1. T è una trasformazione **misurabile** se $T^{-1}A \in \mathcal{B}_1$, per ogni $A \in \mathcal{B}_2$
2. T è una trasformazione che **preserva la misura** se è misurabile e se $m_1(T^{-1}A) = m_2(A)$, per ogni $A \in \mathcal{B}_2$
3. T trasformazione **invertibile** se sia T che T^{-1} preservano la misura.

La composizione di trasformazioni che preservano la misura è ancora una trasformazione che preserva la misura. Cioè le trasformazioni che preservano la misura sono i morfismi tra gli spazi di misura.

Il seguente lemma risulta utile nel caso in cui non si conoscano direttamente le σ -algebre \mathcal{B}_1 e \mathcal{B}_2 :

Lemma 1.1.2. Sia $T : (X_1, \mathcal{B}_1, m_1) \rightarrow (X_2, \mathcal{B}_2, m_2)$ una trasformazione qualsiasi e sia $\mathcal{A} \subseteq \mathcal{B}_2$ un'algebra che genera \mathcal{B} . Se $T^{-1}(\mathcal{A}) \subseteq \mathcal{B}_1$ e $m_1(T^{-1}A) = m_2(A)$, per ogni $A \in \mathcal{A}$, allora T preserva la misura.

Dimostrazione. Sia $\mathcal{C} = \{B \in \mathcal{B}_2 \mid T^{-1}B \in \mathcal{B}_1, m_1(T^{-1}B) = m_2(B)\}$. Allora $\mathcal{A} \subseteq \mathcal{C}$ per ipotesi e inoltre \mathcal{C} è una classe monotona, quindi coincide con la σ -algebra generata da \mathcal{A} . Dunque $\mathcal{C} = \mathcal{B}_2$. \square

Possiamo, ora, dare qualche esempio abbastanza classico di trasformazioni che preservano la misura. Di tali esempi studieremo anche altre proprietà che verranno introdotte successivamente.

Esempio 1.1.3. L'applicazione identità su (X, \mathcal{B}, m) è una trasformazione che preserva la misura.

Lemma 1.1.4 (Esistenza delle misure di Haar).

Dato un gruppo topologico compatto G , allora esiste una misura finita m (detta **misura di Haar**) definita sui boreliani e tale che:

$$m(a \cdot E) = m(E) \quad \forall a \in G, \forall E \subseteq G \text{ boreliano.}$$

Lemma 1.1.5 (Unicità della misura di Haar).

Se m e μ sono entrambe misure finite di Haar su un gruppo topologico compatto G , allora esiste una costante $c > 0$ tale che $m = c \cdot \mu$; quindi esiste un'unica misura di Haar normalizzata.

Esempio 1.1.6. Sia $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ il cerchio unitario dotato della σ -algebra \mathcal{B} dei boreliani e della misura di Haar m . Se $T : S^1 \rightarrow S^1$ è la rotazione $T(z) = az$, con $a \in S^1$, allora T è una trasformazione che preserva la misura (perché m è di Haar).

In effetti vale in generale che ogni rotazione $T(x) = ax$ definita su un gruppo topologico compatto G (con $a \in G$ fissato) preserva la misura di Haar.

Esempio 1.1.7. Ogni isomorfismo (di gruppi) continuo di un gruppo topologico compatto in sé preserva la misura di Haar.

Dimostrazione. Sia $\phi : G \rightarrow G$ un isomorfismo continuo e sia m la misura di Haar su G . Allora $\mu(E) := m(\phi^{-1}(E))$ è una misura di probabilità (sui boreliani) e vale

$$\mu(\phi(x) \cdot E) = m(\phi^{-1}(\phi(x) \cdot E)) = m(x \cdot \phi^{-1}(E)) = \mu(E).$$

Ma siccome ϕ mappa G surgettivamente in sé, allora dall'unicità della misura di Haar segue che $\mu = m$. \square

Esempio 1.1.8. Sia G un gruppo topologico compatto, sia $a \in G$ fissato e sia $A : G \rightarrow G$ un omomorfismo continuo surgettivo. Allora, ogni trasformazione affine della forma $T(x) = a \cdot A(x)$ preserva la misura di Haar perché è composizione di trasformazioni che la preservano.

Esempio 1.1.9 (Trasformazioni shift).

Sia $Y = \{1, \dots, k\}$ e si dia a ogni i una misura p_i in modo tale che $\sum_{i=1}^k p_i = 1$. Si definisca $X = \prod_{-\infty}^{+\infty} Y$ dotato della misura prodotto e la trasformazione $T : X \rightarrow X$ detta (p_1, \dots, p_k) -**shift bilatero** (**unilatero** se $X = \prod_0^{+\infty} Y$) data da

$$T(\dots, x_0, x_1, \dots) = (\dots, x_1, x_2, \dots)$$

Allora T preserva la misura di ogni rettangolo misurabile e quindi T preserva la misura degli insiemi che sono unione disgiunta finita di rettangoli misurabili e quindi (per il lemma precedente) T preserva la misura.

Esempio 1.1.10. Si consideri $I = [0, 1]$ dotato della misura di Lebesgue. Allora le proiezioni $p_1, p_2 : I^2 \rightarrow I$ definite rispettivamente da $p_1(x, y) = x$ e $p_2(x, y) = y$, sono trasformazioni che preservano la misura.

Osservazione 1.1.11. Sia X_1 un insieme qualsiasi, sia $(X_2, \mathcal{B}_2, m_2)$ uno spazio di probabilità e sia $T : X_1 \rightarrow X_2$ una funzione surgettiva. Allora si possono scegliere su X_1 una σ -algebra \mathcal{B}_1 e una misura m_1 tali che T preserva tali misure: infatti basta definire $\mathcal{B}_1 = T^{-1}(\mathcal{B}_2)$ e $m_1(T^{-1}B) = m_2(B)$, per ogni $B \in \mathcal{B}_2$.

Teorema 1.1.12 (di ricorrenza di Poincaré).

Sia T una trasformazione che preserva la misura sullo spazio di probabilità (X, \mathcal{B}, m) . Se $E \in \mathcal{B}$ è tale che $m(E) > 0$, allora quasi tutti i punti di E

sono ricorrenti (rispetto ad E), cioè per quasi ogni $x \in E$ si ha $T^n(x) \in E$ per un opportuno $n = n(x) \in \mathbb{N}$.

Inoltre esiste $F \subseteq E$, con $m(F) = m(E)$, tale che per ogni $x \in F$ esiste una successione crescente di indici interi positivi n_i per i quali $T^{n_i}x \in F$, per ogni $i \in \mathbb{N}$.

Dimostrazione. Sia $N \geq 0$ e sia $E_N = \bigcup_{n=N}^{\infty} T^{-n}E$; allora

$$T^{-1}E_N = T^{-1} \left(\bigcup_{n=N}^{\infty} T^{-n}E \right) = \bigcup_{n=N}^{\infty} T^{-n-1}E = E_{N+1} ;$$

inoltre per costruzione si ha $E_{N+1} \subseteq E_N \subseteq \dots \subseteq E_0$ e per ipotesi su T si ha $m(E_{N+1}) = m(T^{-1}E_N) = m(E_N)$. E quindi per ogni $N \in \mathbb{N}$ si ha $m(E_N) = m(E_0)$ e $m(\bigcap_{N \in \mathbb{N}} E_N) = m(E_0)$. A questo punto osserviamo che

- $\bigcap_{N \in \mathbb{N}} E_N = \bigcap_{N \in \mathbb{N}} (\bigcup_{n=N}^{\infty} T^{-n}E)$ è l'insieme di tutti i punti che cadono in E frequentemente rispetto all'iterazione di T (proprio perché, ad esempio, $T^{-n}E$ è l'insieme dei punti che cadono in E dopo n iterazioni di T);
- $F = E \cap (\bigcap_{N \in \mathbb{N}} E_N)$ consiste di tutti i punti di E che cadono frequentemente in E rispetto all'iterazione di T ;
- infine $\bigcap_{N \in \mathbb{N}} E_N \subseteq E_0$.

Da quanto appena detto si ottiene che

$$m(F) = m(E \cap \bigcap_{N \in \mathbb{N}} E_N) = m(E \cap E_0) = m(E) .$$

Resta da mostrare che i punti di F sono ricorrenti in F : se $x \in F$, fissata $\{n_i\} \subseteq \mathbb{N}$ crescente tale che $T^{n_i}(x) \in E$ per ogni $i \in \mathbb{N}$, allora per ogni $j \in \mathbb{N}$ si ha $T^{n_j}(x) \in F$ poiché $\{T^{n_i - n_j}(T^{n_j}(x))\}_{i \geq j} \subseteq E$. Dunque $\{T^{n_j}(x)\}_{j \in \mathbb{N}} \subseteq F$. \square

Osservazione 1.1.13. Nel teorema precedente non c'è bisogno che T preservi la misura, ma basta che T sia **incomprimibile** cioè che valga $m(B) = m(T^{-1}B)$ per ogni $B \in \mathcal{B}$ tale che $T^{-1}B \subseteq B$.

Definizione 1.1.14. Una trasformazione misurabile $T : X \rightarrow X$ è detta **dissipativa** se esiste $F \subset X$ misurabile e con misura $m(F) > 0$ e tale che gli insiemi $F, T^{-1}F, T^{-2}F, \dots$ sono disgiunti a due a due. Altrimenti T è detta **conservativa**.

Osservazione 1.1.15. Nella dimostrazione della prima parte del teorema di Poincaré 1.1.12 entrambe le ipotesi che T preserva la misura e che $m(X) < +\infty$ vengono usate in un senso debole. Infatti quello che è veramente essenziale è la non esistenza di un insieme $F \subseteq X$ di misura positiva e tale che $F, T^{-1}F, T^{-2}F, \dots$ siano a due a due disgiunti, altrimenti si avrebbe l'assurdo che

$$+\infty > m(X) \geq m \left(\bigcup_{n \in \mathbb{N}} T^{-n}F \right) = \sum_{n \in \mathbb{N}} m(T^{-n}F) = \sum_{n \in \mathbb{N}} m(F) = +\infty .$$

Allora risulta chiaro che la prima parte del teorema 1.1.12 resta valida nel caso in cui T sia una trasformazione conservativa.

Osservazione 1.1.16. La tesi del teorema di Poincaré 1.1.12 può essere riformulata in termini di funzioni indicatrici: per quasi ogni $x \in E$ la serie $\sum_{n \in \mathbb{N}} \chi_E(T^n x)$ diverge. Inoltre vale in generale che per q.o. $x \in \{f > 0\}$ la serie $\sum_n f(T^n x)$ diverge, per ogni funzione $f \geq 0$ misurabile. Infatti se si definisce $E_k = \{f > \frac{1}{k}\}$ per ogni $k > 0$, allora per il teorema di ricorrenza 1.1.12 per q.o. $x \in E_k$ si ha $T^n x \in E_k$ frequentemente. Dunque si ha la tesi considerando $x \in \bigcup_k E_k = \{f > 0\}$ e la maggiorazione con $\sum_k \frac{1}{k} = +\infty$.

1.2 Convergenza in media

Per il teorema di ricorrenza di Poincaré 1.1.12, quasi ogni punto di $E \subseteq X$ ritorna in E infinite volte. Quanto “tempo” trascorrono in E tali punti? In altre parole, dato $x \in E$ e dato $n \in \mathbb{N}$ esiste finito o infinito il limite

$$\lim_{n \rightarrow +\infty} \frac{\text{cardinalità}(\{0 \leq k < n \mid T^k x \in E\})}{n} \quad ?$$

Se $f = \chi_E$ è la funzione indicatrice di un insieme misurabile $E \subseteq X$, allora il limite precedente è uguale a $\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$ e quindi il problema precedente diventa un problema di convergenza secondo Cesàro della successione $\{f(T^k x)\}_{k \in \mathbb{N}}$.

Dato uno spazio di misura (X, \mathcal{B}, m) indicheremo con $L^p(X, \mathcal{B}, m)$ (o semplicemente con L^p qualora non vi sia pericolo di confusione) lo spazio delle (classi di) funzioni da X a valori in \mathbb{C} p -sommabili, cioè per cui valga $\|f\|_p := (\int_X |f|^p dm)^{\frac{1}{p}} < +\infty$. Nel caso particolare di funzioni da X in \mathbb{R} indicheremo tale spazio con $L^p_{\mathbb{R}}(X, \mathcal{B}, m)$.

Osservazione 1.2.1. Sia (X, \mathcal{B}, m) uno spazio di misura. Se $f : X \rightarrow \mathbb{R}$ (oppure \mathbb{C}) e se si definisce $g(x) = f(Tx)$, allora $g = Uf$ definisce un operatore U tra spazi di funzioni con le seguenti proprietà:

1. U è lineare:

$$U(af + bg)(x) = (af + bg)(Tx) = af(Tx) + bg(Tx) = aUf(x) + bUg(x) \\ \forall a, b \in \mathbb{C} \quad \forall f, g : X \rightarrow \mathbb{C}$$

2. Se T preserva la misura allora U è ben definito da $L^1(X, \mathcal{B}, m)$ in sé (ed anzi, U è un'isometria di L^1): la dimostrazione si fa per approssimazioni successive. Se $f = \chi_E$, con $E \in \mathcal{B}$ allora si ha

$$\|f\|_1 = \int_X f dm = m(E) = m(T^{-1}E) = \int_X \chi_{T^{-1}E} dm = \\ = \int_X (\chi_E \circ T) dm = \|Uf\|_1$$

Dalla linearità si ha $\|f\|_1 = \|Uf\|_1$ anche per combinazioni lineari di funzioni indicatrici, cioè per funzioni semplici. Quindi si ha la tesi per $f \geq 0$ applicando il teorema di Beppo Levi ad una successione crescente di funzioni semplici approssimante f ; infine per quanto riguarda il caso con f di segno qualsiasi basta considerare che $\|f\|_1 = \| |f| \|_1$. (Vale la pena notare che in nessuno dei casi precedenti è stato necessario avere $m(X) < +\infty$).

3. U è un'isometria di $L^2(X, \mathcal{B}, m)$ in sé: infatti, lo è di L^1 e quindi

$$\|f\|_2^2 = \int_X f^2 dm = \|f^2\|_1 = \|Uf^2\|_1 = \int_X f^2(Tx) dm = \|Uf\|_2^2$$

4. Se T è invertibile allora U è un'isometria invertibile (quindi è un operatore unitario sullo spazio di Hilbert L^2): infatti l'inverso di U è l'operatore definito da $Vf(x) = f(T^{-1}x)$.

Lemma 1.2.2. *Sia \mathcal{H} uno spazio di Hilbert e sia $U : \mathcal{H} \rightarrow \mathcal{H}$ un'isometria lineare. Allora:*

$$Uf = f \iff U^*f = f$$

dove U^* indica l'operatore aggiunto.

Dimostrazione.

(\Rightarrow) Siccome $Uf = f$, allora $U^*(Uf) = U^*f$; ma essendo U unitario allora $U^*U = id$ e quindi $f = U^*f$. (Vale la pena notare che nel caso in cui U sia l'operatore associato ad una trasformazione $T : X \rightarrow X$ allora la dimostrazione è la stessa sia per $m(X) < +\infty$ che per $m(X) = +\infty$. Tuttavia nel secondo caso non è detto che valga l'implicazione $U^*U = id \Rightarrow UU^* = id$. Si veda l'esempio 1.2.3).

(\Leftarrow) Se $U^*f = f$ allora $\|Uf - f\| = 0$, infatti:

$$\|Uf - f\|^2 = \langle Uf - f, Uf - f \rangle = \|Uf\|^2 - \langle f, Uf \rangle - \langle Uf, f \rangle + \|f\|^2,$$

ma per definizione di operatore aggiunto per ogni $f, g \in \mathcal{H}$ si ha $\langle f, Ug \rangle = \langle U^*f, g \rangle$, e quindi vale che $\langle Uf, f \rangle = \langle f, U^*f \rangle = \|f\|^2$ (dove l'ultima uguaglianza sussiste in quanto U è unitario per ipotesi) e analogamente $\langle f, Uf \rangle = \langle U^*f, f \rangle = \|f\|^2$, e da cui la tesi. \square

Esempio 1.2.3. Sia $X = [0, +\infty[$ e sia $Uf(x) = f(x-1)\chi_{[1, +\infty[}(x)$. Allora U è un operatore associato alla trasformazione

$$Tx = \begin{cases} x-1 & , \text{ se } x \geq 1 \\ 0 & , \text{ se } x \in [0, 1[\end{cases}$$

Si può verificare che $\|Uf\|_{L^2(0, +\infty)}^2 = \|f\|_{L^2(0, +\infty)}^2$ e che l'operatore aggiunto è definito da $U^*g(x) = g(x+1)$. Infine, si ha che $U^*U = id$, ma $UU^* \neq id$ poiché

$$\begin{aligned} U^*Uf(x) &= (Uf)(x+1) = f(x)\chi_{[1, \infty[}(x+1) = f(x) \quad \text{e} \\ UU^*f(x) &= g(x)\chi_{[1, \infty[}(x) \end{aligned}$$

Teorema 1.2.4 (della media ergodica).

Se U è un operatore unitario su uno spazio di Hilbert complesso \mathcal{H} e se P è la proiezione sul sottospazio dei vettori U -invarianti, allora $\frac{1}{n} \sum_{j=0}^{n-1} U^j f$ converge (in norma) a Pf , per ogni $f \in \mathcal{H}$.

Dimostrazione. Si procede in vari passi.

1. Se $Uf = f$, cioè f è nel sottospazio dei vettori U -invarianti, allora $f = Pf$ e ovviamente $\frac{1}{n} \sum_{j=0}^{n-1} U^j f = f$ converge a Pf .

2. Se $f = g - Ug$ per qualche $g \in \mathcal{H}$, allora $\sum_{j=0}^{n-1} U^j f = g - U^n g$ perché

$$\sum_{j=0}^{n-1} U^j f = \sum_{j=0}^{n-1} U^j (g - Ug) = \sum_{j=0}^{n-1} U^j g - \sum_{j=1}^n U^j g = U^0 g - U^n g = g - U^n g. \quad (1.1)$$

Da ciò segue che

$$0 \leq \left\| \frac{1}{n} \sum_{j=0}^{n-1} U^j f \right\| \leq \frac{1}{n} \|g - U^n g\| \leq \frac{2}{n} \|g\| \xrightarrow{n \rightarrow +\infty} 0 \quad (1.2)$$

3. L'insieme $\mathcal{E} = \{f \in \mathcal{H} \mid \exists g \in \mathcal{H} : f = g - Ug\}$ è un sottospazio vettoriale (non necessariamente chiuso) di \mathcal{H} . Siccome l'operatore $A_n = \frac{1}{n} \sum_{j=0}^{n-1} U^j$ è uniformemente limitato (perché U è unitario), allora da ciò segue che $A_n f \xrightarrow{n} 0$ per ogni $f \in \mathcal{E}$. In generale, se $f_k \xrightarrow{k} f$ e se $A_n f_k \xrightarrow{n} 0$ per ogni $k \in \mathbb{N}$, allora $A_n f \xrightarrow{n} 0$. Infatti, ciò segue dal fatto che prendendo $\|f - f_k\| < \varepsilon$ sufficientemente piccolo si ha $\|A_n(f - f_k)\| < \delta$ ed anche $\|A_n f_k\| < \delta$, quindi:

$$\|A_n f\| \leq \|A_n(f - f_k)\| + \|A_n f_k\| < 2\delta.$$

Siccome $(\overline{\mathcal{E}})^\perp = \mathcal{E}^\perp$ (dove \perp indica il sottospazio ortogonale), allora per ogni $h \in \mathcal{E}^\perp$ si ha $\langle h, g - Ug \rangle = 0$ per ogni $g \in \mathcal{H}$, quindi

$$0 = \langle h, g \rangle - \langle h, Ug \rangle = \langle h, g \rangle - \langle U^* h, g \rangle,$$

ovvero, per ogni $h \in \mathcal{E}^\perp$

$$\langle h - U^* h, g \rangle = 0 \quad \forall g \in \mathcal{H}.$$

Da ciò segue che $h - U^* h = 0$, cioè $h = U^* h$ e quindi per il lemma 1.2.2 $h = Uh$. Viceversa, se $Uh = h$ allora $\langle h, Ug - g \rangle = 0$ per ogni g , infatti (usando il lemma 1.2.2)

$$\langle h, Ug - g \rangle = \langle h, Ug \rangle - \langle h, g \rangle = \langle U^* h, g \rangle - \langle h, g \rangle = \langle U^* h, g \rangle - \langle U^* h, g \rangle = 0.$$

In conclusione, $\mathcal{E}^\perp = \{f \in \mathcal{H} \mid Uf = f\}$ è il sottospazio dei vettori U -invarianti.

4. Dal passo (3) si ottiene che per ogni $f \in \mathcal{H} = \mathcal{E} \oplus \mathcal{E}^\perp$ esistono $f_1 = g - Ug \in \mathcal{E}$ ed $f_2 = Uf_2 \in \mathcal{E}^\perp$ tali che $f = f_1 + f_2$. La tesi allora segue applicando ad f_1 ed f_2 i risultati ottenuti rispettivamente nel punto (2) ed (1).

□

1.3 Convergenza puntuale e teoremi ergodici

Teorema 1.3.1 (ergodico massimale).

Sia $U : L_{\mathbb{R}}^1(X, \mathcal{B}, m) \rightarrow L_{\mathbb{R}}^1(X, \mathcal{B}, m)$ un operatore lineare e positivo (cioè

$f \geq 0 \Rightarrow Uf \geq 0$) con norma $\|U\| \leq 1$ e sia $N > 0$ intero. Sia $f \in L^1_{\mathbb{R}}(X, \mathcal{B}, m)$ e si definiscano

$$\begin{cases} f_0 = 0 \\ f_n = f + Uf + \dots + U^{n-1}f \end{cases} \quad e \quad F_N = \max_{0 \leq n \leq N} f_n .$$

Allora

$$\int_{\{x \in X \mid F_N(x) > 0\}} f \, dm \geq 0 .$$

Dimostrazione. Chiaramente $F_N \in L^1_{\mathbb{R}}(X, \mathcal{B}, m)$ e $F_N \geq f_n$ per ogni $n \leq N$, quindi per la positività di U si ha che $UF_N \geq Uf_n$ per ogni $n \leq N$. Quindi si ha

$$UF_N + f \geq Uf_n + f = f_n + 1 \quad \forall n \leq N ,$$

da cui

$$UF_N + f \geq \max_{1 \leq n \leq N+1} f_n \geq \max_{1 \leq n \leq N} f_n = \max_{0 \leq n \leq N} f_n = F_N .$$

Dunque vale $f \geq F_N - UF_N$ sull'insieme $A = \{x \in X \mid F_N(x) > 0\}$, da cui la tesi, poiché

$$\int_A f \, dm \geq \int_A F_N \, dm - \int_A UF_N \, dm \geq 0 ,$$

in quanto

$$\begin{aligned} \int_A UF_N \, dm &= \int_A \left[U(F_N \chi_A) - U(|F_N| \chi_{A^c}) \right] dm \leq \int_A U(F_N \chi_A) \, dm \leq \\ &\leq \int_X |U(F_N \chi_A)| \, dm \leq \int_X |F_N \chi_A| \, dm = \int_A F_N \, dm . \end{aligned}$$

□

Osservazione 1.3.2. Naturalmente, le ipotesi del teorema ergodico massimale 1.3.1 sono verificate nel caso in cui l'operatore U sia indotto da una trasformazione T che preserva la misura.

Corollario 1.3.3. Sia $T : X \rightarrow X$ una trasformazione che preserva la misura e sia $g \in L^1_{\mathbb{R}}(X, \mathcal{B}, m)$. Inoltre sia

$$B_\alpha = \left\{ x \in X \mid \sup_{n \geq 1} \left[\frac{1}{n} \sum_{k=0}^{n-1} g(T^k x) \right] > \alpha \right\} .$$

Se $A \subseteq X$ ha misura finita ed è T -invariante (cioè tale che $m(A) < +\infty$ e $T^{-1}A = A$), allora

$$\int_{B_\alpha \cap A} g \, dm \geq \alpha \cdot m(B_\alpha \cap A) .$$

Dimostrazione.

- Innanzitutto consideriamo il caso in cui $m(X) < +\infty$ e $A = X$. Sia $f = g - \alpha$, allora $B_\alpha = \bigcup_{N \in \mathbb{N}} \{x \in X \mid F_N(x) > 0\}$ (dove le F_N sono definite come nel teorema 1.3.1) e quindi $\int_{B_\alpha} f \, dm \geq 0$ per il teorema 1.3.1. Dunque $\int_{B_\alpha} g \, dm \geq \alpha \cdot m(B_\alpha)$.

- In generale, basta usare la restrizione $T|_A$ invece di T .

□

Teorema 1.3.4 (ergodico individuale di Birkhoff).

Sia $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ una trasformazione che preserva la misura (con $m(X) \leq +\infty$) e sia $f \in L^1(X, \mathcal{B}, m)$. Allora:

1. esiste $f^* \in L^1(X, \mathcal{B}, m)$ tale che $\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i \xrightarrow[n \rightarrow +\infty]{} f^*$ q.o.;
2. $f^* \circ T = f^*$, cioè f^* è T -invariante;
3. se $m(X) < +\infty$ allora $\int_X f \, dm = \int_X f^* \, dm$.

Dimostrazione. È sufficiente mostrare il teorema per $f \in L^1_{\mathbb{R}}$. Si definiscano

$$f^*(x) := \limsup_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \quad \text{e} \quad f_*(x) := \liminf_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x).$$

Allora sia f^* che f_* sono T -invarianti, infatti posto $a_n(x) := \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$ allora

$$\frac{n+1}{n} a_n(x) - a_n(Tx) = \frac{n+1}{n} \frac{1}{n+1} \sum_{i=0}^n f(T^i x) - \frac{1}{n} \sum_{i=1}^n f(T^i x) = \frac{f(x)}{n},$$

e dunque:

$$f^* \circ T(x) = \limsup_n a_n(Tx) = \limsup_n \left[\frac{n+1}{n} a_{n+1}(x) - \frac{f(x)}{n} \right] = \limsup_n a_{n+1} = f^*(x).$$

Analogamente $f_* \circ T = f_*$. Ora, si definisca per $\beta < \alpha$

$$E_{\alpha, \beta} := \{x \in X \mid f_*(x) < \beta, \alpha < f^*(x)\},$$

allora si ha che

$$T^{-1} E_{\alpha, \beta} = E_{\alpha, \beta} \quad \text{e} \quad E_{\alpha, \beta} \cap \left\{ x \in X \mid \sup_{n \geq 1} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) > \alpha \right\} = E_{\alpha, \beta}.$$

A questo punto, si dimostra che $m(E_{\alpha, \beta}) < +\infty$ in modo da poter applicare il corollario 1.3.3:

- nel caso in cui $\alpha > 0$, sia $C \subseteq E_{\alpha, \beta}$ tale che $m(C) < +\infty$. Allora $h := f - \alpha \cdot \chi_C$ è una funzione integrabile e per il teorema massimale ergodico 1.3.1 vale che:

$$\int_{\bigcup_{N \in \mathbb{N}} \{x \in X \mid H_N(x) > 0\}} (f - \alpha \cdot \chi_C) \, dm \geq 0, \quad (1.3)$$

dove H_N è definita come nel teorema massimale ergodico 1.3.1. Tuttavia, siccome $C \subseteq \bigcup_{N \in \mathbb{N}} \{x \in X \mid H_N(x) > 0\}$, allora si ottiene:

$$\int_X |f| \, dm \geq \int_{\bigcup_{N \in \mathbb{N}} \{H_N > 0\}} f \, dm \geq \alpha \int_{\bigcup_N \{H_N > 0\}} \chi_C \, dm = \alpha \cdot m(C).$$

Dunque $m(C) \leq \frac{1}{\alpha} \int_X |f| \, dm$ per ogni $C \subseteq E_{\alpha, \beta}$ tale che $m(C) < +\infty$, quindi $m(E_{\alpha, \beta}) < +\infty$.

- nel caso in cui $\alpha \leq 0$, allora deve essere anche $\beta < 0$. Quindi applicando il ragionamento precedente sostituendo $-f$ ad f e $-\beta$ ad α si ottiene che $m(E_{\alpha,\beta}) < +\infty$.

Sia ora

$$B_\alpha = \left\{ x \in X \mid \sup_{n \geq 1} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) > \alpha \right\};$$

per quanto visto sopra $B_\alpha \subseteq E_{\alpha,\beta}$. Dal corollario 1.3.3 segue che

$$\int_{E_{\alpha,\beta}} f \, dm = \int_{E_{\alpha,\beta} \cap B_\alpha} f \, dm \geq \alpha \cdot m(E_{\alpha,\beta} \cap B_\alpha) = \alpha \cdot m(E_{\alpha,\beta})$$

e dunque:

$$\int_{E_{\alpha,\beta}} f \, dm \geq \alpha \cdot m(E_{\alpha,\beta}).$$

Ora, sostituendo f, α, β rispettivamente con $-f, -\alpha, -\beta$ si ottiene:

$$(-f)^* = -f_* \quad , \quad (-f)_* = -f^* \quad \text{e} \quad \int_{E_{\alpha,\beta}} f \, dm \leq \beta \cdot m(E_{\alpha,\beta}).$$

Dunque si è ottenuto che

$$\alpha \cdot m(E_{\alpha,\beta}) \leq \int_{E_{\alpha,\beta}} f \, dm \leq \beta \cdot m(E_{\alpha,\beta})$$

ed, essendo $\beta < \alpha$, si ha $m(E_{\alpha,\beta}) = 0$.

Inoltre, siccome

$$\{x \in X \mid f_*(x) < f^*(x)\} \subseteq \bigcup_{\{\alpha,\beta \in \mathbb{Q} \mid \beta < \alpha\}} E_{\alpha,\beta},$$

allora si ha $m(\{x \in X \mid f_*(x) < f^*(x)\}) = 0$, ovvero:

$$f^*(x) = f_*(x) \quad \text{q.o..}$$

Per dimostrare che $f_* \in L^1$ basta usare il Lemma di Fatou, infatti posto $g_n(x) = \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \right|$ si ha:

$$\begin{aligned} \int_X g_n \, dm &= \int_X \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \right| \, dm = \frac{1}{n} \sum_{i=0}^{n-1} \int_X |f(T^i x)| \, dm = \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \int_X |U^i f| \, dm \leq \frac{1}{n} \sum_{i=0}^{n-1} \int_X |f| \, dm = \int_X |f| \, dm; \end{aligned}$$

quindi $\liminf_n \int g_n < +\infty$ e per il Lemma di Fatou $\liminf_n g_n = |f_*|$ è integrabile. Dunque f_* è integrabile e quindi $f^* = f_* \in L^1$. Questo completa la dimostrazione della (1) e della (2).

Si assuma, ora, che $m(X) < +\infty$ e si definisca:

$$D_k^n = \left\{ x \in X \mid \frac{k}{n} \leq f^*(x) < \frac{k+1}{n} \right\} \quad \text{dove} \quad k \in \mathbb{Z} \quad \text{e} \quad n \in \mathbb{N} \setminus \{0\}.$$

Per ogni $\varepsilon > 0$ si ha $D_k^n \cap B_{(\frac{k}{n} - \varepsilon)} = D_k^n$, e quindi dal corollario 1.3.3 segue:

$$\int_{D_k^n} f \, dm \geq \left(\frac{k}{n} - \varepsilon \right) m(D_k^n) \quad \forall \varepsilon > 0,$$

da cui

$$\int_{D_k^n} f \, dm \geq \frac{k}{n} m(D_k^n). \quad (1.4)$$

Allora per la (1.4) si ha:

$$\int_{D_k^n} f^* \, dm \leq \frac{k+1}{n} m(D_k^n) \leq \frac{1}{n} m(D_k^n) + \int_{D_k^n} f \, dm$$

e sommando su $k \in \mathbb{Z}$ si ottiene

$$\int_X f^* \, dm \leq \frac{m(X)}{n} + \int_X f \, dm \quad \forall n \in \mathbb{N} \setminus \{0\},$$

dunque $\int_X f^* \, dm \leq \int_X f \, dm$. Con lo stesso argomento, sostituendo $-f$ a f , si ottiene

$$\int_X (-f)^* \, dm \leq \int_X (-f) \, dm \quad \text{cioè} \quad - \int_X f_* \, dm \leq - \int_X f \, dm,$$

ed essendo $f^* = f_*$ q.o., si ottiene la tesi (3). \square

Corollario 1.3.5. *Sia $m(X) < +\infty$, sia $T : X \rightarrow X$ una trasformazione che preserva la misura e sia $f \in L^1$. Allora, posto $f^*(x) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$, si ha:*

$$\lim_{m \rightarrow \infty} \int_X \left| \frac{1}{m} \sum_{i=0}^{m-1} f(T^i x) - f^*(x) \right| = 0.$$

Overo, nell'ipotesi di $m(X) < +\infty$, la convergenza è in L^1 (e non solo q.o. come nel teorema 1.3.4).

Dimostrazione. Vi sono due casi:

1. Se f è limitata, allora anche le medie $\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$ sono limitate (dallo stesso limite) e quindi la tesi segue applicando il teorema della convergenza dominata e il teorema ergodico individuale di Birkhoff 1.3.4:

$$\lim_n \int_X \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) - f^*(x) \right| dm = \int_X \lim_n \left| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) - f^*(x) \right| dm = 0.$$

2. Se f non è limitata, si può usare un'approssimazione con una funzione $g \in L^1$ limitata e tale che siano abbastanza piccoli i due numeri (usando anche la parte (1) appena dimostrata, con la scelta di $n \in \mathbb{N}$ sufficientemente grande):

$$\|f - g\|_1 \quad \text{e} \quad \left\| \frac{1}{n} \sum_{i=0}^{n-1} g \circ T^i - g^* \right\|_1$$

A questo punto, la tesi segue osservando che

$$\begin{aligned} \left\| \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i - f^* \right\|_1 &\leq \left\| \frac{1}{n} \sum_i (f \circ T^i - g \circ T^i) \right\|_1 + \\ &+ \left\| \frac{1}{n} \sum_i g \circ T^i - g^* \right\|_1 + \|g^* - f^*\|_1. \end{aligned}$$

e che

$$\left\| \frac{1}{n} \sum_i (f \circ T^i - g \circ T^i) \right\|_1 \leq \|f - g\|_1 \quad \text{e}$$

$$\|g^* - f^*\|_1 = \|g - f\|_1 \quad (\text{per il teorema individuale ergodico 1.3.4}).$$

□

Definizione 1.3.6. Un operatore lineare $S : L^1(X, m) \rightarrow L^1(X, m)$ si dice **doppiamente stocastico** se per ogni $f \in L^1$ soddisfa le proprietà:

1. $f \geq 0 \implies Sf \geq 0$
2. $\int_X Sf \, dm = \int_X f \, dm$
3. $S1 = 1$ (q.o.), dove $1 : X \rightarrow \mathbb{C} \quad x \mapsto 1$

Proposizione 1.3.7. Se T è doppiamente stocastico su uno spazio di misura finita X , allora mappa L^p in sé per ogni $p \in [1, +\infty]$, con norme che soddisfano $\|T\|_{\mathcal{L}(L^p)} \leq 1$ e $\|T\|_{\mathcal{L}(L^1)} = \|T\|_{\mathcal{L}(L^\infty)} = 1$. Inoltre, anche T^* è doppiamente stocastico.

Dimostrazione. Innanzitutto, siccome $m(X) < +\infty$ (e senza perdita di generalità si può assumere $m(X)=1$) allora $L^\infty \subseteq \dots \subseteq L^p \subseteq \dots \subseteq L^1$ per ogni $p \in (1, +\infty)$, e quindi T è ben definito su tutti gli spazi in questione. Per ogni $f \in L^\infty$, siano $f^+, f^- \geq 0$ la parte positiva e negativa di f , in modo che:

$$f = f^+ - f^- \quad \text{e} \quad |f| = f^+ + f^- .$$

Essendo T doppiamente stocastico, anche $Tf^+, Tf^- \geq 0$ e quindi:

$$|Tf| = |T(f^+ - f^-)| = |Tf^+ - Tf^-| \leq Tf^+ + Tf^- = T|f|. \quad (1.5)$$

Inoltre, per le proprietà (1) e (3) della definizione 1.3.6 si ha:

$$T|f| \leq T\|f\|_\infty = \|f\|_\infty. \quad (1.6)$$

Combinando la (1.5) e la (1.6) si ottiene $\|Tf\|_\infty \leq \|f\|_\infty$, ovvero $\|T\|_{\mathcal{L}(L^\infty)} \leq 1$. Ma siccome $T1 = 1$ per definizione, allora $\|T\|_{\mathcal{L}(L^\infty)} = 1$. (Vale la pena notare che per quanto dimostrato finora è stato necessario usare solo le proprietà (1) e (3) della definizione 1.3.6.) La proprietà (2) della 1.3.6 equivale a dire che $\langle Tf, 1 \rangle = \langle f, T^*1 \rangle = \langle f, 1 \rangle$ per ogni $f \in L^1$, e quindi è equivalente a $T^*1 = 1$. Segue, quindi, che $\|T^*\|_{\mathcal{L}(L^\infty)} = \|T\|_{\mathcal{L}(L^1)} = 1$. Per quanto riguarda la disuguaglianza $\|T\|_p \leq 1$ con $p \in (1, \infty)$, essa segue da un'applicazione del teorema di convessità di Riesz.

Resta da mostrare che T^* mappa L^1 in sé e che soddisfa la condizione (2) della 1.3.6: se $f \in L^\infty$ allora

$$\int_X T^* f \, dm = \langle 1, T^* f \rangle = \langle T1, f \rangle = \langle 1, f \rangle = \int_X f \, dm.$$

Ora, siccome L^∞ è denso in L^1 , allora T^* ha un'unica estensione continua a tutto L^1 che soddisfa

$$\int_X T^* f \, dm = \int_X f \, dm \quad \forall f \in L^1.$$

Inoltre, integrando la (1.5) sostituendo T^* a T si ottiene:

$$\int_X |T^* f| \, dm \leq \int_X T^* |f| \, dm = \int_X |f| \, dm < +\infty.$$

□

Si può ora generalizzare il teorema della media ergodica 1.2.4, ottenibile dal teorema seguente prendendo $p = 2$ in modo da ottenere lo spazio di Hilbert $\mathcal{H} = L^2$.

Teorema 1.3.8 (della media ergodica di Yosida).

Sia U un operatore doppiamente stocastico sullo spazio di probabilità X (cioè $m(X) = 1$) e sia $f \in L^p$. Allora esiste $f^* \in L^p$ tale che:

$$\lim_{n \rightarrow \infty} \left\| \frac{1}{n} \sum_{i=0}^{n-1} U^i f - f^* \right\|_p = 0 \quad (1.7)$$

Dimostrazione. Si può procedere per vari passi.

- Sia $h : X \rightarrow \mathbb{R}$ tale che $Uh \leq h$. Posto $g := h - Uh$, si ha $g \geq 0$ e per la proprietà (2) di 1.3.6

$$\int_X g \, dm = 0,$$

quindi $g = 0$ q.o., cioè

$$Uh = h \quad \text{q.o.}$$

Per la proposizione 1.3.7, quanto appena visto vale anche per U^* ; in particolare, se $U^*h_1 = h_1$ e $U^*h_2 = h_2$ allora, posto $(h_1 \wedge h_2)(x) := \min_{x \in X}(h_1(x), h_2(x))$, si ha che $U^*(h_1 \wedge h_2) = h_1 \wedge h_2$.

- Sia ora $f = g - Ug$ per una opportuna $g \in L^p$. Ponendo

$$U_n f(x) = \sum_{i=0}^{n-1} U^i f(x),$$

si ottiene $U_n f = g - U^n g$ ed anche (in maniera analoga a quanto fatto in 1.1 e 1.2 nella dimostrazione del teorema della media ergodica 1.2.4):

$$\left\| \frac{1}{n} U_n f \right\|_p \leq \frac{2}{n} \|g\|_p \rightarrow_{n \rightarrow \infty} 0.$$

Quindi la tesi 1.7 vale se $f^* = 0$ e se

$$f \in \mathcal{H}_1 := \{g - Ug \mid g \in L^p\}.$$

Facilmente si verifica che la tesi vale anche se $f^* = f$ e se

$$f \in \mathcal{H}_2 := \{g \in L^p \mid Ug = g\}.$$

perché

$$\left\| \frac{1}{n} \sum_{i=0}^{n-1} U^i f - f^* \right\|_p = \left\| \frac{1}{n} \sum_{i=0}^{n-1} f - f \right\|_p = \|f - f\|_p = 0$$

- Procediamo ora mostrando che $\mathcal{H}_1 + \mathcal{H}_2$ è denso in L^p e a tal fine basta mostrare che l'unica funzione F ortogonale a \mathcal{H}_1 e ad \mathcal{H}_2 è la funzione nulla (q.o.). Infatti:

– siccome $F \perp \mathcal{H}_1$, allora

$$\forall g \in L^p \quad \langle F, g - Ug \rangle = \langle F, g \rangle - \langle U^*F, g \rangle = 0,$$

quindi $U^*F = F$. Ora, per ogni $c \in \mathbb{R}$ fissato, si definisca $A_c := \{x \in X \mid F(x) > c\}$ e per ogni $\varepsilon > 0$ sia

$$g_\varepsilon := \frac{1}{\varepsilon} [(c + \varepsilon) \wedge F - (c \wedge F)].$$

Si noti che $0 \leq g_\varepsilon \leq 1$ e che $g_\varepsilon \nearrow \chi_{A_c}$ (se $\varepsilon \rightarrow 0$), quindi $U^*g_\varepsilon \nearrow U^*\chi_{A_c}$. Inoltre, siccome le funzioni $F, c, c + \varepsilon$ sono tutte U^* -invarianti, si ha

$$U^*g_\varepsilon = \frac{1}{\varepsilon} [U^*((c + \varepsilon) \wedge F) - U^*(c \wedge F)] = \frac{1}{\varepsilon} [(c + \varepsilon) \wedge F - (c \wedge F)] = g_\varepsilon,$$

quindi $U^*\chi_{A_c} = \lim_\varepsilon U^*g_\varepsilon = \lim g_\varepsilon = \chi_{A_c}$. In realtà vale anche che $U\chi_{A_c} = \chi_{A_c}$, infatti dato $B \in \mathcal{B}$ si ha

$$U^*\chi_{A_c \cap B} \leq U^*\chi_{A_c} = \chi_{A_c} \quad \text{e} \quad U^*\chi_{A_c \cap B} \leq U^*\chi_B$$

quindi

$$U^*\chi_{A_c \cap B} = \chi_{A_c} \cdot U^*\chi_{A_c \cap B} \leq \chi_{A_c} \cdot U^*\chi_B. \quad (1.8)$$

Analogamente $U^*\chi_{A_c \cap B} \leq \chi_{A_c} \cdot U^*\chi_B$, ovvero valgono

$$\begin{aligned} U^*((1 - \chi_{A_c})\chi_B) &\leq (1 - \chi_{A_c}) \cdot U^*\chi_B = U^*\chi_B - U^*\chi_{A_c \cap B} \leq \\ &\leq U^*\chi_B - \chi_{A_c} \cdot U^*\chi_B \end{aligned} \quad (1.9)$$

$$\chi_{A_c} \cdot U^*\chi_B \leq U^*\chi_{A_c \cap B}.$$

Usando (1.8) e (1.9) si ottiene $U^*\chi_{A_c \cap B} = \chi_{A_c} \cdot U^*\chi_B$, dunque

$$\begin{aligned} \langle U\chi_{A_c}, \chi_B \rangle &= \langle \chi_{A_c}, U^*\chi_B \rangle = \langle 1, \chi_{A_c} U^*\chi_B \rangle = \langle 1, U^*\chi_{A_c \cap B} \rangle = \\ &= \langle U1, \chi_{A_c \cap B} \rangle = \langle 1, \chi_{A_c \cap B} \rangle = \langle \chi_{A_c}, \chi_B \rangle, \end{aligned}$$

e per l'arbitrarietà di B si ottiene $U\chi_{A_c} = \chi_{A_c}$.

- Ora, siccome $m(X) < +\infty$ allora $\chi_A \in L^\infty \subset L^p$ e quindi $\chi_A \in \mathcal{H}_2$. Siccome $F \perp \mathcal{H}_2$, allora vale

$$0 = \langle \chi_A, F \rangle = \int_{\{F > c\}} F \, dm \quad \forall c \in \mathbb{R},$$

quindi $F = 0$ q.o. Dunque $\mathcal{H}_1 + \mathcal{H}_2$ è denso in L^p .

- Ora, per densità, data $f \in L^p$ esistono $f_k \in \mathcal{H}_1 + \mathcal{H}_2$ per cui vale la (1.7) (per quanto visto sopra) e tali che $f_k \rightarrow f$ in L^p . Allora

$$\begin{aligned} \left\| \frac{1}{n} U_n f - \frac{1}{m} U_m f \right\|_p &\leq \frac{1}{n} \sum_{i=0}^{n-1} \|U^i(f - f_k)\|_p + \frac{1}{m} \sum_{i=0}^{m-1} \|U^i(f - f_k)\|_p + \\ &+ \left\| \frac{1}{n} U_n f_k - \frac{1}{m} U_m f_k \right\|_p \leq 2\|f - f_k\|_p + \left\| \frac{1}{n} U_n f_k - \frac{1}{m} U_m f_k \right\|_p \end{aligned}$$

da cui segue la tesi. □

1.4 Ergodicità

Se T è una trasformazione che preserva la misura su X e se $X = E \cup F$ con

$$m(E), m(F) > 0, \quad E \cap F = \emptyset, \quad TE = E \quad \text{e} \quad TF = F,$$

allora lo studio delle proprietà di T su X si può ridurre allo studio (presumibilmente più semplice) delle proprietà di $T|_E$ e $T|_F$. Un sistema verrà detto ergodico quando questa scomposizione non può essere fatta.

Definizione 1.4.1. Una trasformazione T dello spazio di misura (X, \mathcal{B}, m) è detta **ergodica** se per ogni $B \in \mathcal{B}$ tale che $T^{-1}B = B$ si ha:

$$m(B) = 0 \quad \text{o} \quad m(X \setminus B) = 0.$$

Proposizione 1.4.2. T è ergodica se e solo se ogni funzione $f : X \rightarrow \mathbb{R}$ misurabile e T -invariante è costante (q.o.).

Dimostrazione.

(\Leftarrow) Sia $A \subseteq X$ un invariante: allora la sua funzione indicatrice χ_A è invariante, quindi $\chi_A = \text{costante}$ (q.o.), ossia $m(A) = 0$ oppure $m(X \setminus A) = 0$.

(\Rightarrow) Sia f misurabile e T -invariante; si definisca:

$$X(k, n) = \left\{ x \in X \left| \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \right. \right\} \quad \forall k \in \mathbb{Z}, \forall n \in \mathbb{N}.$$

Siccome f è T -invariante, allora anche gli insiemi $X(k, n)$ sono T -invarianti e quindi per ipotesi per ogni fissato $n \in \mathbb{N}$ tutti tranne uno degli $X(k, n)$ (indicizzato da $k = k_n$) hanno misura 0. A questo punto, osserviamo che per ogni $n \in \mathbb{N}$ vale $X(k_{n+1}, n+1) \subseteq X(k_n, n)$, quindi esiste $\lim_n k_n/2^n =: \bar{t} \in \mathbb{R}$. Inoltre, vale

$$m(\{x \in X | f(x) \neq \bar{t}\}) = \lim_{n \rightarrow \infty} m(X \setminus X(k_n, n)) = \lim_n 0 = 0.$$

Quindi $f(x) = \bar{t}$ q.o. □

Osservazione 1.4.3. Nel caso in cui $m(X) < +\infty$, è vero anche che T è ergodica se e solo se ogni funzione T -invariante in L^1 (o in L^2) è costante (q.o.). La dimostrazione di questo fatto è banale in quanto ogni funzione indicatrice $f = \chi_A$ è misurabile se e solo se è in L^1 (poiché $m(X) < +\infty$) ovvero se e solo se è in L^2 .

Teorema 1.4.4. *Se $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ preserva la misura e $m(X) = 1$, allora sono fatti equivalenti:*

1. T è ergodica;
2. Se $B \in \mathcal{B}$ e $m(T^{-1}B \Delta B) = 0$ allora $m(B) = 0$ o $m(B) = 1$ (dove Δ indica la differenza simmetrica tra insiemi);
3. Per ogni $A, B \in \mathcal{B}$ con $m(A), m(B) > 0$ esiste $n \in \mathbb{N} \setminus \{0\}$ tale che $m(T^{-n}A \cap B) > 0$.

Dimostrazione.

(1 \Rightarrow 2) Sia $B \in \mathcal{B}$ tale che $m(T^{-1}B \Delta B) = 0$ e sia

$$\mathcal{B} \ni B_\infty := \bigcap_{n \in \mathbb{N}} \bigcup_{i=n}^{\infty} T^{-i}B$$

allora $T^{-1}B_\infty = B_\infty$ e $m(B_\infty) = m(B)$.

Quindi, dall'ipotesi, $m(B_\infty) = 0$ o $m(B_\infty) = 1$, cioè $m(B) = 0$ o $m(B) = 1$.

(2 \Rightarrow 3) Siano $A, B \in \mathcal{B}$ tali che $m(A), m(B) > 0$ e si supponga per assurdo che (3) sia falsa, cioè che

$$\forall n \in \mathbb{N} \setminus \{0\} \quad m(T^{-n}A \cap B) = 0.$$

Dunque

$$m\left(\left(\bigcup_{n=1}^{+\infty} T^{-n}A\right) \cap B\right) = 0.$$

Posto $A' = \bigcup_{n=1}^{+\infty} T^{-n}A$, allora

$$T^{-1}A' \subseteq A' \quad \text{e} \quad m(A') = m(T^{-1}A'),$$

quindi

$$m(T^{-1}A' \Delta A') = 0$$

e per ipotesi

$$m(A') = 0 \quad \text{oppure} \quad m(A') = 1.$$

Ma siccome $T^{-1}A \subseteq A'$ e T preserva la misura, allora $m(A') \geq m(A) > 0$, quindi $m(A') = 1$. Ma questo contraddice il fatto che $m(A' \cap B) = 0$, perché $m(X) = 1$ e $m(B) > 0$.

(3 \Rightarrow 1) Per assurdo esiste $B \in \mathcal{B}$ tale che $T^{-1}B = B$ e $0 < m(B) < 1$. Allora

$$m(T^{-n}B \cap B^c) = 0 \quad \forall n > 0,$$

cosa che contraddice la (3). □

Osservazione 1.4.5. Sia X uno spazio metrico compatto di probabilità con misura di Borel m positiva su ogni aperto non vuoto di X , e sia $T : X \rightarrow X$ continua ed ergodica (rispetto a m). Allora

$$m\left(\left\{x \in X \mid \{T^n x \mid n \geq 0\} \text{ è denso in } X\right\}\right) = 1.$$

Ovvero, quasi ogni punto di X ha orbita $\{T^n x\}_n$ densa.

Dimostrazione. Sia $\{U_n\}_{n \in \mathbb{N}}$ una base per la topologia di X . Allora:

$$\{T^n x \mid n \geq 0\} \text{ è denso in } X \iff x \in \bigcap_{n \in \mathbb{N}} \bigcup_{k \in \mathbb{N}} T^{-k} U_n.$$

Siccome

$$T^{-1} \left(\bigcup_{k \in \mathbb{N}} T^{-k} U_n \right) \subseteq \bigcup_{k \in \mathbb{N}} T^{-k} U_n$$

e T è ergodica e preserva la misura, allora si ha che

$$m \left(\bigcup_{k \in \mathbb{N}} T^{-k} U_n \right) = 0 \quad \text{oppure} \quad 1.$$

Ma siccome $\bigcup_{k \in \mathbb{N}} T^{-k} U_n$ è un aperto non vuoto, allora per ipotesi ha misura positiva, quindi 1, da cui la tesi. \square

Esempio 1.4.6. La trasformazione identica *id* sullo spazio di probabilità (X, \mathcal{B}, m) è ergodica se e solo se tutti gli elementi di \mathcal{B} hanno misura 0 o 1.

Esempio 1.4.7. Sia $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ la circonferenza unitaria e sia $T : S^1 \rightarrow S^1$ definita da $T(z) = az$, con $a \in \mathbb{C}$ qualsiasi. Allora:

$$T \text{ è ergodica} \iff a \text{ non è radice dell'unità.}$$

Dimostrazione.

(\Rightarrow) Sia a radice dell'unità, cioè $a^p = 1$ per un opportuno $p \in \mathbb{N} \setminus \{0\}$.

Sia $f(z) := z^p$, chiaramente $f \neq \text{costante}$ (q.o.) ed è T -invariante perché $(f \circ T)(z) = f(az) = a^p z^p = z^p = f(z)$, quindi T non può essere ergodica.

(\Leftarrow) Se a non è radice dell'unità, sia $f \circ T = f$ con $f \in L^2$ e se ne consideri lo sviluppo in serie $f(z) = \sum_{n \in \mathbb{Z}} b_n z^n$.

Allora

$$\sum_n b_n a^n z^n = (f \circ T)(z) = f(z) = \sum_n b_n z^n$$

e quindi $b_n(a^n - 1) = 0$ per ogni $n \in \mathbb{N}$; in particolare, per $n \neq 0$ si ha $b_n = 0$, cioè $f = \text{costante}$ (q.o.). La tesi segue dall'osservazione 1.4.3. \square

Esempio 1.4.8. Sullo spazio discreto \mathbb{N} con la misura data dalla cardinalità $m(A) = \#(A)$, dove $A \subseteq \mathbb{N}$, la traslazione $T_1 x = x + 1$ è ergodica, mentre la traslazione $T_2 x = x + 2$ non lo è (infatti, ad esempio, l'insieme dei numeri pari è invariante, ma né esso né il suo complementare hanno misura nulla).

Esempio 1.4.9. Il (p_1, \dots, p_k) -shift bilatero è ergodico. (Analogamente quello unilatero).

Dimostrazione. Sia \mathcal{A} l'algebra generata dalle unioni finite di rettangoli misurabili. Sia $E \in \mathcal{B}$ un insieme T -invariante e sia $\varepsilon > 0$ fissato, allora esiste $A \in \mathcal{A}$ tale che $m(E \Delta A) < \varepsilon$, quindi

$$\begin{aligned} |m(E) - m(A)| &= |m(E \cap A) + m(E \setminus A) - m(A \cap E) - m(A \setminus E)| < \\ &< m(E \setminus A) + m(A \setminus E) < \varepsilon. \end{aligned}$$

A questo punto, siccome A è nell'algebra \mathcal{A} allora si può prendere n abbastanza grande in modo che $B := T^{-n}A$ non dipenda dalle stesse coordinate di A . In questo modo si ha anche

$$\begin{aligned} m(A \cap B) &= m(A)m(B) = m(A)^2 \\ m(E \Delta B) &= m(E \Delta A) < \varepsilon \end{aligned}$$

Ma siccome $E \Delta (A \cap B) \subset (E \Delta A) \cup (E \Delta B)$, allora $m(E \Delta (A \cap B)) < 2\varepsilon$ e quindi $|m(E) - m(A \cap B)| < 2\varepsilon$. Infine (siccome $m(A), m(E) \leq 1$)

$$\begin{aligned} |m(E) - m(E)^2| &\leq |m(E) - m(A \cap B)| + |m(A \cap B) - m(E)^2| < \\ &< 2\varepsilon + |m(A)^2 - m(E)^2| \leq \\ &\leq 2\varepsilon + m(A) |m(A) - m(E)| + m(E) |m(A) - m(E)| < 4\varepsilon. \end{aligned}$$

E dunque, per l'arbitrarietà di ε si ha che $m(E) = m(E)^2$ e quindi che $m(E) = 0$ oppure 1. Un argomento analogo si applica anche nel caso di shift unilatero. \square

Esempio 1.4.10. La traslazione $Tx = x + 1$, definita su \mathbb{R} dotato della misura di Lebesgue, non è ergodica. Infatti un esempio di insieme invariante e ma con misura non banale è

$$\bigcup_{n \in \mathbb{N}} \left] n, n + \frac{1}{2} \right[.$$

Osservazione 1.4.11. Nello spazio di Hilbert $L^2(S^1, \mathcal{B}, m)$, dove m è la misura di Haar costituita dalla misura di Lebesgue normalizzata e \mathcal{B} è la σ -algebra dei Boreliani, le funzioni $f_n(x) := x^n$ con $n \in \mathbb{Z}$ formano una base ortonormale.

Osservazione 1.4.12 (Alcune proprietà dell'insieme dei caratteri).

Sia G un gruppo topologico abeliano localmente compatto e con una base numerabile (si ricordi che per un gruppo topologico compatto si ha: G metrico se e soltanto se G a base numerabile) e sia $\widehat{G} = \{\gamma : G \rightarrow S^1 \mid \gamma \text{ omeomorfismo continuo}\}$ l'insieme dei caratteri di G , che è un gruppo abeliano con l'operazione di moltiplicazione (puntuale) di funzioni.

Inoltre, \widehat{G} è un gruppo abeliano localmente compatto a base numerabile con la topologia compatto-aperto, definita in generale tra due spazi topologici Y, Z dalla pre-base:

$$\bigcup_{\substack{K \subset Y \text{ compatto;} \\ U \subset Z \text{ aperto}}} \{f : Y \rightarrow Z \mid f(K) \subseteq U\}.$$

Inoltre valgono le seguenti proprietà:

1. G compatto $\iff \widehat{G}$ discreto.

2. \widehat{G} è naturalmente isomorfo (come gruppo topologico) a G , tramite la mappa:

$$\widehat{G} \rightarrow G \quad \alpha \mapsto a,$$

dove $\alpha(\gamma) = \gamma(a) \quad \forall \gamma \in \widehat{G}$.

3. $\widehat{G_1 \times G_2} \simeq \widehat{G_1} \times \widehat{G_2}$, dove \simeq e \times indicano l'isomorfismo e il prodotto diretto di gruppi.
4. Se $H \subsetneq G$ è un sottogruppo chiuso, allora esiste $\gamma \in \widehat{G}$ con $\gamma \neq 1$ tale che $\gamma(h) = 1 \quad \forall h \in H$.
5. Se G è compatto allora gli elementi di \widehat{G} sono mutuamente ortogonali in $L^2(m)$, dove m è una misura di Haar.
Infatti basta mostrare che

$$\int_G \gamma(x) dm(x) = 0 \quad \forall \gamma \neq 1.$$

Ma se $a \in G$, essendo m di Haar e γ un omeomorfismo, allora

$$\int \gamma(x) dm(x) = \int \gamma(ax) dm(x) = \gamma(a) \cdot \int \gamma dm,$$

da cui segue la tesi prendendo $a \in G$ tale che $\gamma(a) \neq 1$.

6. Se G è compatto allora gli elementi di \widehat{G} formano una base ortonormale di $L^2(m)$, dove m è la misura di Haar normalizzata.
7. Se $A : G \rightarrow G$ è un endomorfismo di G (cioè un omomorfismo continuo), si può definire un endomorfismo di \widehat{G} come segue:

$$\widehat{A} : \widehat{G} \rightarrow \widehat{G} \quad \widehat{A}(\gamma) = \gamma \circ A.$$

E valgono:

$$\begin{aligned} A \text{ iniettiva} &\iff \widehat{A} \text{ surgettiva} \\ A \text{ surgettiva} &\iff \widehat{A} \text{ iniettiva.} \end{aligned}$$

Quindi:

$$A \text{ automorfismo} \iff \widehat{A} \text{ automorfismo}$$

Esempio 1.4.13. La rotazione $T(x) = ax$ di un gruppo metrico compatto G con $m(G) = 1$ è ergodica se e soltanto se l'insieme $\{a^n\}_{n \in \mathbb{Z}}$ è denso in G .
(In particolare: se T ergodica allora G è abeliano.)

Dimostrazione.

(\Rightarrow) Se $Tx = ax$ è ergodica, allora dall'osservazione 1.4.5 segue che $\{a^n x_0\}_{n \in \mathbb{Z}}$ è denso in G per qualche $x_0 \in G$ e quindi $\{a^n\}_n$ è denso (infatti: se $y \in G$ allora $\exists \{n_i\}_i \subset \mathbb{N}$ tale che $a^{n_i} x_0 \rightarrow_i y x_0$, e quindi $a^{n_i} \rightarrow_i y$).

(\Leftarrow) Sia $\{a^n\}_{n \in \mathbb{Z}} \subset G$ denso. Ciò implica che G è abeliano, infatti se $a^{n_h} \rightarrow_h b$ e se $a^{n_k} \rightarrow_k c$, allora

$$a^{n_h} a^{n_k} \xrightarrow{h,k} bc$$

$$a^{n_k} a^{n_h} \xrightarrow{h,k} cb,$$

quindi $bc = cb$ per ogni $b, c \in G$.

Ora, sia $f \in L^2$ e T -invariante e sia $\widehat{G} = \{\gamma : G \rightarrow S^1 \mid \gamma \text{ omeomorfismo continuo}\}$ l'insieme dei caratteri di G , che forma una base ortonormale di $L^2(G)$ perché G è compatto. Allora si può scrivere

$$f = \sum_i b_i \gamma_i \quad \text{con } \gamma_i \in \widehat{G}$$

e allora per l'invarianza di f

$$\sum_i b_i \gamma_i(a) \gamma_i(x) = \sum_i b_i \gamma_i(x),$$

così che se $b_i \neq 0$ allora $\gamma_i(a) = 1$, da cui, per densità, $\gamma_i \equiv 1$. Dunque il solo termine costante della serie di f può essere diverso da 0, cioè f è costante (q.o.). \square

Proposizione 1.4.14. *Sia (X, \mathcal{B}, m) uno spazio di probabilità e sia $T : X \rightarrow X$ una trasformazione che preserva la misura. Allora:*

$$T \text{ è ergodica} \iff \forall A, B \in \mathcal{B} \quad \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A \cap B) \xrightarrow{n \rightarrow \infty} m(A)m(B).$$

Dimostrazione.

(\Rightarrow) Prendendo $f := \chi_A$ nel teorema di Birkhoff 1.3.4 si ottiene:

$$\frac{1}{n} \sum_{i=0}^{n-1} \chi_A \circ T^i \xrightarrow{n \rightarrow \infty} m(A) \quad \text{q.o.}$$

Moltiplicando per χ_B :

$$\frac{1}{n} \sum_{i=0}^{n-1} (\chi_A \circ T^i) \cdot \chi_B \xrightarrow{n \rightarrow \infty} m(A) \cdot \chi_B \quad \text{q.o.}$$

Per il teorema di convergenza dominata, integrando si ottiene:

$$\frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A \cap B) \xrightarrow{n \rightarrow \infty} m(A)m(B).$$

(\Leftarrow) Se $T^{-1}E = E$ con $E \in \mathcal{B}$, presi $A = B = E$ allora dall'ipotesi

$$\frac{1}{n} \sum_{i=0}^{n-1} m(E) \xrightarrow{n \rightarrow \infty} m(E)^2.$$

Quindi $m(E) = m(E)^2$, così che $m(E) = 0$ oppure 1. \square

Teorema 1.4.15 (ergodico di Von Neumann).

Sia T una trasformazione che preserva la misura sullo spazio di probabilità (X, \mathcal{B}, m) . Se $f \in L^p(m)$ (con $1 \leq p < +\infty$), allora esiste $f^ \in L^p(m)$ tale che:*

$$f^* \circ T = f^*$$

$$\lim_{n \rightarrow \infty} \left\| \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) - f^*(x) \right\|_p = 0$$

Dimostrazione. Se g è limitata e misurabile, allora $g \in L^\infty \subseteq L^p$ (perché $m(X) < +\infty$) e quindi per il teorema ergodico

$$\frac{1}{n} \sum_{i=0}^{n-1} g(T^i x) \xrightarrow[n \rightarrow \infty]{} g^*(x) \quad (q.o.)$$

E chiaramente anche g^* è limitata, cioè $g^* \in L^\infty \subseteq L^p$.

Inoltre

$$\left| \frac{1}{n} \sum_i g(T^i x) - g^*(x) \right|^p \rightarrow_n 0 \quad (q.o.)$$

e quindi per il teorema di convergenza dominata

$$\left\| \frac{1}{n} \sum_i g(T^i x) - g^*(x) \right\|_p \rightarrow_n 0.$$

Cioè per ogni $\varepsilon > 0$, esiste $N = N(\varepsilon, g) \in \mathbb{N}$ tale che:

$$\forall n > N, \forall k > 0 \quad \left\| \frac{1}{n} \sum_{i=0}^{n-1} g(T^i x) - \frac{1}{n+k} \sum_{i=0}^{n+k-1} g(T^i x) \right\|_p < \frac{\varepsilon}{2}.$$

Ora, per $f \in L^p$ si definisca

$$M_n(f)(x) := \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x).$$

La successione $\{M_n f\}_n$ è di Cauchy in L^p , infatti $\|M_n f\|_p \leq \|f\|_p$ e scelta $g \in L^\infty$ con $\|f - g\|_p < \varepsilon/4$ (cosa possibile per densità di L^∞ in L^p) allora:

$$\begin{aligned} & \|M_n f - M_{n+k} f\|_p \leq \\ & \leq \|M_n f - M_n g\|_p + \|M_n g + M_{n+k} g\|_p + \|M_{n+k} g - M_{n+k} f\|_p \leq \\ & \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Poiché $M_n g \rightarrow g^*$ in L^p per ogni $g \in L^\infty$, per densità si ottiene $M_n f \rightarrow f^*$ in L^p per ogni $f \in L^p$. D'altronde, essendo

$$\left(\frac{n+1}{n} \right) M_{n+1} f(x) - M_n f(Tx) = \frac{f(x)}{n},$$

per $n \rightarrow \infty$ segue che $f^* \circ T = f^*$ (q.o.). □

Osservazione 1.4.16. Se T è una trasformazione che preserva la misura ed è ergodica e se $f \in L^1$, allora dal teorema precedente segue che $f^* = \lim_n \frac{1}{n} \sum_i (f \circ T^i)$ è invariante, quindi:

- se $m(X) < +\infty$ allora $f^* = \text{costante}$ (q.o.), e inoltre dal teorema ergodico di Birkhoff 1.3.4 siccome $\int_X f \, dm = \int_X f^* \, dm$ allora

$$f^* \equiv \frac{1}{m(X)} \int_X f \, dm.$$

Cioè, in spazi di misura finita la media spaziale $\frac{1}{m(X)} \int f$ è uguale (q.o.) alla media temporale (asintotica) $f^* = \lim_n \frac{1}{n} \sum_i f \circ T^i$ per ogni funzione $f \in L^1$.

- se $m(X) = +\infty$ allora $f^* = \text{costante}$ (q.o.), ma allora necessariamente $f^* = 0$ (q.o.).

Esempio 1.4.17 (Teorema di Borel sui **numeri normali**).

Quasi tutti i numeri nell'intervallo $[0, 1)$ sono normali in base 2, cioè per q.o. $x \in [0, 1)$, detto k_n il numero di cifre 1 nelle prime n cifre dell'espressione binaria di x , si ha

$$\lim_{n \rightarrow \infty} \frac{k_n}{n} = \frac{1}{2}.$$

Dimostrazione. Sia $T : [0, 1) \rightarrow [0, 1)$ definita da $T(x) = 2x \pmod{1}$, allora

- T è una trasformazione che preserva la misura di Lebesgue, e basta verificarlo per gli aperti $(a, b) \subset [0, 1)$:

$$\begin{aligned} m(T^{-1}(a, b)) &= m\left(\left(\frac{a}{2}, \frac{b}{2}\right) \cup \left(\frac{a}{2} + \frac{1}{2}, \frac{b}{2} + \frac{1}{2}\right)\right) = \\ &= \left(\frac{b}{2} - \frac{a}{2}\right) + \left(\frac{b}{2} + \frac{1}{2} - \frac{a}{2} - \frac{1}{2}\right) = b - a = m(a, b). \end{aligned}$$

- T è ergodica, infatti sia $f : [0, 1) \rightarrow \mathbb{C}$ con $f \in L^2$ (si veda l'osservazione 1.4.3) e T -invariante, allora se si considera lo sviluppo in serie di $f(x) = \sum_{n \in \mathbb{Z}} a_n x^n$ si ottiene:

$$\sum_{n \in \mathbb{Z}} a_n x^n = f(x) = f(2x \pmod{1}) = \sum_{n \in \mathbb{Z}} a_n 2^n x^n,$$

da cui

$$2^n a_n = a_n \quad \forall n \in \mathbb{Z},$$

cioè $a_n = 0$ per ogni $n \neq 0$. Dunque $f = \text{costante}$ (q.o.).

Ora, si consideri un'espansione binaria (si ricordi che in generale tale espansione non è unica, ma i punti che hanno due espansioni diverse formano un insieme di misura nulla):

$$x = \frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \dots,$$

allora

$$T(x) = T\left(\frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \dots\right) = \frac{a_2}{2^2} + \frac{a_3}{2^3} + \dots$$

Posto $f(x) = \chi_{[1/2, 1)}(x)$, allora

$$f(T^i x) = f\left(\frac{a_{i+1}}{2} + \frac{a_{i+2}}{2^2} + \dots\right) = \begin{cases} 1 & \text{sse } a_{i+1} = 1 \\ 0 & \text{sse } a_{i+1} = 0 \end{cases}.$$

Dunque il numero di 1 nelle prime n cifre dell'espansione binaria di x è dato da $\sum_{i=0}^{n-1} f(T^i(x))$, quindi applicando il teorema ergodico di Birkhoff 1.3.4 e l'osservazione precedente si ottiene (e usando il fatto che l'insieme dei razionali ha misura nulla):

$$\frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) \xrightarrow{n \rightarrow \infty} f^*(x) \equiv \frac{1}{m(X)} \int_X f \, dm = \int_X \chi_{[1/2,1)} \, dm = \frac{1}{2} \quad (\text{q.o.}).$$

□

Osservazione 1.4.18. In uno spazio di misura finita X , una trasformazione $T : X \rightarrow X$ è ergodica se e solo se la media temporale di ogni funzione integrabile è costante. Basta provare che ogni funzione invariante e integrabile è costante, ma ciò segue dal fatto che una funzione invariante coincide con la sua stessa media temporale. Per spazi di misura infinita quanto appena visto non vale, ad esempio basta prendere $X = \mathbb{Z} \times \mathbb{Z}$ e $T(x, y) = (x + 1, y)$ e vedere che T non è ergodico e la costante 0 è la sola funzione invariante.

La proposizione seguente risponde alla domanda opposta rispetto ai teoremi visti finora: cosa si può dire sull'integrabilità di f conoscendo f^* ?

Proposizione 1.4.19. *Siano T una trasformazione ergodica e che preserva la misura, $m(X) < +\infty$ e $f \geq 0$ una funzione qualsiasi su X . Se $\frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$ ha limite finito per q.o. $x \in X$, allora f è integrabile.*

Dimostrazione. Sia $f^*(x) := \lim_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$ (definita ad esempio ponendo $f^*(x) = 0$ per gli $x \in X$ per cui il limite non è finito o non esiste, e tale insieme di punti x ha misura nulla). Allora tale funzione è T -invariante e integrabile, quindi per l'ergodicità di T essa è costante (q.o.).

Ora, sia f_k la funzione ottenuta da f troncata all'altezza k quando tale soglia viene superata, allora f_k è limitata e quindi integrabile. Per il teorema di convergenza monotona $\int f_k \, dm \rightarrow_k \int f \, dm$, e quindi basta mostrare la limitatezza della successione $\{\int f_k \, dm\}_k$ per ottenere la tesi. Ma infatti, siccome $0 \leq f_k \leq f$, allora

$$\frac{1}{n} \sum_{i=0}^{n-1} f_k(T^i x) \leq \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$$

e quindi passando al limite per $n \rightarrow +\infty$ si ottiene $f_k^* \leq f^* \equiv c$ (q.o.).

Dunque

$$\int_X f_k^* \, dm \leq c \cdot m(X) \quad \forall k \in \mathbb{N},$$

e per il teorema ergodico di Birkhoff 1.3.4

$$\int_X f_k^* \, dm = \int_X f_k \, dm$$

quindi la successione $\{\int f_k \, dm\}_k$ è limitata. □

1.5 Mixing e altre proprietà

Definizione 1.5.1. Sia (X, \mathcal{B}, m) uno spazio di probabilità e sia $T : X \rightarrow X$ una trasformazione che preserva la misura. Allora si definisce:

1. T **ergodica** sse $\frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A \cap B) \xrightarrow{n \rightarrow \infty} m(A)m(B)$
2. T **weak-mixing** sse $\frac{1}{n} \sum_{i=0}^{n-1} |m(T^{-i}A \cap B) - m(A)m(B)| \xrightarrow{n \rightarrow \infty} 0$
3. T **strong-mixing** sse $m(T^{-n}A \cap B) \xrightarrow{n \rightarrow \infty} m(A)m(B)$.

(Si veda anche la proposizione 1.4.14).

Osservazione 1.5.2. T strong-mixing $\implies T$ weak-mixing $\implies T$ ergodica. Questo segue dal fatto generale che se $\{a_n\}_{n \in \mathbb{N}}$ è una successione di numeri reali qualsiasi, allora:

$$a_n \xrightarrow{n \rightarrow \infty} 0 \implies \frac{1}{n} \sum_{i=0}^{n-1} |a_i| \xrightarrow{n \rightarrow \infty} 0 \implies \frac{1}{n} \sum_{i=0}^{n-1} a_i \xrightarrow{n \rightarrow \infty} 0.$$

Lemma 1.5.3. Sia $T : X \rightarrow X$ trasformazione che preserva la misura e sia \mathcal{A} un'algebra che genera la σ -algebra \mathcal{B} . Se le condizioni di ergodicità/weak-mixing/strong-mixing sono verificate per tutti gli elementi di \mathcal{A} , allora valgono per tutti gli elementi di \mathcal{B} .

Dimostrazione. Sia $\varepsilon > 0$ e siano $E, F \in \mathcal{B}$. Presi $E_0, F_0 \in \mathcal{A}$ tali che $m(E \Delta E_0) < \varepsilon$ e $m(F \Delta F_0) < \varepsilon$ allora

$$\forall n \in \mathbb{N} \quad m\left((T^{-n}E \cap F) \Delta (T^{-n}E_0 \cap F_0)\right) < 2\varepsilon,$$

quindi

$$\forall n \in \mathbb{N} \quad \left| m(T^{-n}E \cap F) - m(T^{-n}E_0 \cap F_0) \right| < 2\varepsilon.$$

Dunque $\forall n \in \mathbb{N}$ si ha:

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}E \cap F) - m(E)m(F) \right| \leq \\ & \leq \left| \frac{1}{n} \sum_{i=0}^{n-1} \left(m(T^{-i}E \cap F) - m(T^{-i}E_0 \cap F_0) \right) \right| + \left| \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}E_0 \cap F_0) - m(E_0)m(F_0) \right| + \\ & + \left| \frac{1}{n} \sum_{i=0}^{n-1} m(E_0)m(F_0) - m(E_0)m(F) \right| + \left| \frac{1}{n} \sum_{i=0}^{n-1} m(E_0)m(F_0) - m(E)m(F) \right| \leq \\ & \leq 2\varepsilon + \left| \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}E_0 \cap F_0) - m(E_0)m(F_0) \right| + \varepsilon + \varepsilon \end{aligned}$$

Allora per l'ergodicità basta osservare che essa vale per E_0, F_0 .

Per la condizione di weak-mixing, portando prima fuori il simbolo di sommatoria si mostra che

$$\left| m(T^{-i}E \cap F) - m(E)m(F) \right| \leq 4\varepsilon + \left| m(T^{-i}E_0 \cap F_0) - m(E_0)m(F_0) \right|,$$

a questo punto basta prendere la media secondo Cesaro su entrambi i membri. Ragionamento analogo si può fare per la condizione di strong-mixing. \square

Esempio 1.5.4. Il (p_1, \dots, p_k) -shift (unilatero o bilatero) è strong-mixing. Infatti, se A e B sono due rettangoli misurabili allora vale

$$m(T^{-n}A \cap B) \xrightarrow[n \rightarrow \infty]{} m(A)m(B),$$

inoltre tale condizione vale per le unioni disgiunte finite di rettangoli misurabili e quindi la tesi segue applicando il lemma precedente.

Lemma 1.5.5. Se $\{a_n\}_{n \in \mathbb{N}}$ è una successione limitata di numeri reali, allora sono equivalenti:

1. $\frac{1}{n} \sum_{i=0}^{n-1} |a_i| \xrightarrow[n \rightarrow \infty]{} 0$

2. esiste $J \subset \mathbb{N}$ con densità

$$\lim_n \frac{\text{cardinalità}(J \cap \{0, 1, \dots, n-1\})}{n} = 0, \text{ tale che } \lim_{n \notin J, n \rightarrow \infty} a_n = 0.$$

3. $\frac{1}{n} \sum_{i=0}^{n-1} |a_i|^2 \xrightarrow[n \rightarrow \infty]{} 0.$

Dimostrazione. Si indichi con $a_M(n) := \text{cardinalità}(M \cap \{0, 1, \dots, n-1\})$, dove $M \subseteq \mathbb{N}$.

(1 \Rightarrow 2) Sia $J_k := \{n \in \mathbb{N} \mid |a_n| \geq 1/k\}$ dove $k > 0$. Allora $J_1 \subset J_2 \subset \dots$; inoltre ogni J_k ha densità 0 poiché

$$\frac{1}{n} \sum_{i=0}^{n-1} |a_i| \geq \frac{1}{n} \frac{1}{k} a_{J_k}(n)$$

ed il primo membro tende a 0 per ipotesi (per $n \rightarrow \infty$). Quindi esistono $0 = l_0 < l_1 < \dots \in \mathbb{N}$ tali che

$$\forall n \geq l_k \quad \frac{1}{n} a_{J_{k+1}}(n) < \frac{1}{k+1}.$$

Ora, sia $J := \bigcup_{k \in \mathbb{N}} (J_{k+1} \cap [l_k, l_{k+1}))$; si può mostrare che J ha densità 0, infatti: innanzitutto siccome $J_1 \subset J_2 \subset \dots$, se $l_k \leq n < l_{k+1}$ si ha

$$J \cap [0, n) = (J \cap [0, l_k)) \cup (J \cap [l_k, n)) \subset (J_k \cap [0, l_k)) \cup (J_{k+1} \cap [0, n))$$

e quindi

$$\frac{1}{n} a_J(n) \leq \frac{1}{n} (a_{J_k}(l_k) + a_{J_{k+1}}(n)) \leq \frac{1}{n} (a_{J_k}(n) + a_{J_{k+1}}(n)) < \frac{1}{k} + \frac{1}{k+1}.$$

Infine $\frac{1}{n} a_J(n) \rightarrow_n 0$, cioè J ha densità 0.

Se $n > l_k$ e $n \notin J$ allora $n \notin J_{k+1}$ e quindi $|a_n| < \frac{1}{k+1}$, dunque $\lim_{n \notin J} |a_n| = 0$.

(2 \Rightarrow 1) Siccome $\{a_n\}_n$ è una successione limitata per ipotesi, sia $|a_n| \leq K$ per ogni $n \in \mathbb{N}$, e infine sia $\varepsilon > 0$ fissato. Allora esiste $N_\varepsilon \in \mathbb{N}$ tale che

$$\forall n \geq N_\varepsilon \quad (n \notin J \implies |a_n| < \varepsilon)$$

ed esiste $M_\varepsilon \in \mathbb{N}$ tale che

$$\forall n \geq M_\varepsilon \quad \frac{a_J(n)}{n} < \varepsilon.$$

Ma allora per ogni $n \geq \max(N_\varepsilon, M_\varepsilon)$ si ha:

$$\frac{1}{n} \sum_{i=0}^{n-1} |a_i| = \frac{1}{n} \left(\sum_{\{0, \dots, n-1\} \cap J} |a_i| + \sum_{\{0, \dots, n-1\} \setminus J} |a_i| \right) < \frac{K}{n} a_J(n) + \varepsilon < (K+1)\varepsilon.$$

(1 \Leftrightarrow 3) Per quanto appena visto nei punti precedenti, basta notare che:

$$\lim_{n \notin J, n \rightarrow \infty} |a_n| = 0 \quad \Leftrightarrow \quad \lim_{n \notin J, n \rightarrow \infty} |a_n|^2 = 0.$$

□

Corollario 1.5.6. *Sono equivalenti:*

1. T weak-mixing;
2. Per ogni $A, B \in \mathcal{B}$ esiste $J = J(A, B) \subset \mathbb{N}$ di densità 0 tale che:

$$\lim_{n \notin J, n \rightarrow \infty} m(T^{-n}A \cap B) = m(A)m(B)$$

3. $\forall A, B \in \mathcal{B} \quad \frac{1}{n} \sum_{i=0}^{n-1} \left| m(T^{-i}A \cap B) - m(A)m(B) \right|^2 \xrightarrow{n \rightarrow \infty} 0.$

Dimostrazione.

Segue dalla definizione e dal lemma precedente. □

Teorema 1.5.7. *Sia (X, \mathcal{B}, m) uno spazio di probabilità e sia $T : X \rightarrow X$ una trasformazione che preserva la misura, con operatore associato $U : L^2 \rightarrow L^2$. Allora valgono:*

1. T ergodica $\Leftrightarrow \forall f, g \in L^2 \quad \frac{1}{n} \sum_{i=0}^{n-1} \langle U^i f, g \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle 1, g \rangle$
 $\Leftrightarrow \forall f \in L^2 \quad \frac{1}{n} \sum_{i=0}^{n-1} \langle U^i f, f \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle 1, f \rangle$
2. T weak-mixing $\Leftrightarrow \forall f, g \in L^2 \quad \frac{1}{n} \sum_{i=0}^{n-1} |\langle U^i f, g \rangle - \langle f, 1 \rangle \langle 1, g \rangle| \xrightarrow{n \rightarrow \infty} 0$
 $\Leftrightarrow \forall f \in L^2 \quad \frac{1}{n} \sum_{i=0}^{n-1} |\langle U^i f, f \rangle - \langle f, 1 \rangle \langle 1, f \rangle| \xrightarrow{n \rightarrow \infty} 0$
3. T strong-mixing $\Leftrightarrow \forall f, g \in L^2 \quad \langle U^n f, g \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle 1, g \rangle$
 $\Leftrightarrow \forall f \in L^2 \quad \langle U^n f, f \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle 1, f \rangle$

Dimostrazione. Per semplicità, si può verificare solo il terzo enunciato in quanto le dimostrazioni sono tutte simili.

Si supponga valida la prima condizione dell'enunciato, cioè T strong-mixing; si può procedere per approssimazioni successive per dimostrare che da ciò segue la terza condizione:

- siano $A, B \in \mathcal{B}$, allora:

$$\int_X \chi_{(T^{-n}A \cap B)} dm = \int_X \chi_A(T^n x) \cdot \chi_B(x) dm \xrightarrow{n \rightarrow \infty} \int_X \chi_A dm \int_X \chi_B dm,$$

cioè $m(T^{-n}A \cap B) \rightarrow_n m(A)m(B)$.

- fissando $B \in \mathcal{B}$, per linearità si ha che per ogni h funzione semplice (cioè combinazione lineare di funzioni indicatrici) vale $\langle U^n h, \chi_B \rangle \rightarrow_n \langle h, 1 \rangle \langle 1, \chi_B \rangle$.
- fissando h si ottiene: $\langle U^n h, h \rangle \rightarrow_n \langle h, 1 \rangle \langle 1, h \rangle$.
- sia $f \in L^2$ qualsiasi e sia $\varepsilon > 0$. Per densità delle funzioni semplici in L^2 si può scegliere h funzione semplice tale che $\|f - h\|_2 < \varepsilon$, inoltre esiste $N = N(\varepsilon) \in \mathbb{N}$ tale che

$$\forall n \geq N \quad |\langle U^n h, h \rangle - \langle h, 1 \rangle \langle 1, h \rangle| < \varepsilon.$$

Allora $\forall n \geq N$ si ha:

$$\begin{aligned} |\langle U^n f, f \rangle - \langle f, 1 \rangle \langle 1, f \rangle| &\leq |\langle U^n f, f \rangle - \langle U^n h, f \rangle| + |\langle U^n h, f \rangle - \langle U^n h, h \rangle| + \\ &+ |\langle U^n h, h \rangle - \langle h, 1 \rangle \langle 1, h \rangle| + |\langle h, 1 \rangle \langle 1, h \rangle - \langle f, 1 \rangle \langle 1, h \rangle| + \\ &+ |\langle f, 1 \rangle \langle 1, h \rangle - \langle f, 1 \rangle \langle 1, f \rangle| \leq \\ &\leq |\langle U^n(f - h), f \rangle| + |\langle U^n h, f - h \rangle| + \varepsilon + \\ &+ |\langle 1, h \rangle| |\langle h - f, 1 \rangle| + |\langle f, 1 \rangle| |\langle 1, h - f \rangle| \leq \\ &\leq \|f - h\|_2 \|f\|_2 + \|f - h\|_2 \|h\|_2 + \varepsilon + \|h\|_2 \|f - h\|_2 + \|f\|_2 \|h - f\|_2 \leq \\ &\leq \varepsilon \|f\|_2 + \|h\|_2 \varepsilon + \varepsilon + \|h\|_2 \varepsilon + \|f\|_2 \varepsilon. \end{aligned}$$

Quindi vale la terza condizione a partire dalla prima.

Ora, supponendo valida la seconda condizione, allora scegliendo $f = \chi_A, g = \chi_B$ dove $A, B \in \mathcal{B}$ si ottiene che T è strong-mixing, cioè la validità della prima condizione.

Infine, supponendo valida la terza condizione, sia $f \in L^2$ e si indichi con \mathcal{H}_f il più piccolo sottospazio chiuso di L^2 contenente f e le funzioni costanti e tale che $U(\mathcal{H}_f) \subseteq \mathcal{H}_f$. Allora

$$\mathcal{F}_f := \left\{ g \in L^2 \mid \langle U^n f, g \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle 1, g \rangle \right\}$$

è anch'esso un sottospazio di L^2 chiuso, contenente f e le costanti e U -invariante; dunque $\mathcal{H}_f \subseteq \mathcal{F}_f$.

Se $g \perp \mathcal{H}_f$ (cioè $g \in \mathcal{H}_f^\perp$) allora per ogni $n \in \mathbb{N}$ si ha $\langle U^n f, g \rangle = 0$ ed anche $\langle 1, g \rangle = 0$ (perché $1 \in \mathcal{H}_f$); quindi $\mathcal{H}_f^\perp \subseteq \mathcal{F}_f$. Ma allora deve essere necessariamente $\mathcal{F}_f = L^2$, cosa che equivale alla seconda condizione. \square

Definizione 1.5.8. Se $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ è una trasformazione che preserva la misura, si definisce la **trasformazione prodotto** $T \times T : X \times X \rightarrow X \times X$ come $(T \times T)(x, y) := (Tx, Ty)$ ed è una trasformazione che preserva la misura sullo **spazio di misura prodotto** $(X \times X, \mathcal{B} \times \mathcal{B}, m \times m)$.

Teorema 1.5.9. *Se T preserva la misura sullo spazio di probabilità (X, \mathcal{B}, m) , allora sono equivalenti:*

1. T weak-mixing,
2. $T \times T$ ergodica,
3. $T \times T$ weak-mixing.

Dimostrazione.

(1 \Rightarrow 3) Siano $A, B, C, D \in \mathcal{B}$, allora (per il corollario 1.5.6) esistono $J_1, J_2 \subset \mathbb{N}$ di densità zero e tali che:

$$\lim_{n \notin J_1, n \rightarrow \infty} m(T^{-n}A \cap B) = m(A)m(B), \quad \lim_{n \notin J_2, n \rightarrow \infty} m(T^{-n}C \cap D) = m(C)m(D).$$

Allora:

$$\begin{aligned} \lim_{n \notin J_1 \cup J_2, n \rightarrow \infty} (m \times m)\left((T \times T)^{-n}(A \times C) \cap (B \times D)\right) &= \\ &= \lim_{n \notin J_1 \cup J_2, n \rightarrow \infty} m(T^{-n}A \cap B)m(T^{-n}C \cap D) = \\ &= m(A)m(B)m(C)m(D) = (m \times m)(A \times C)(m \times m)(B \times D). \end{aligned}$$

Quindi la proprietà di weak-mixing è valida sui rettangoli misurabili della forma $E \times F \in \mathcal{B} \times \mathcal{B}$, quindi è valida anche per unioni finite disgiunte di tali rettangoli, ma questi ultimi formano un'algebra \mathcal{F} che genera la σ -algebra \mathcal{B} .

Allora per il corollario 1.5.6 si ha:

$$\frac{1}{n} \sum_{i=0}^{n-1} |m(T^{-i}A \cap B) - m(A)m(B)| \xrightarrow[n \rightarrow \infty]{} 0 \quad \forall A, B \in \mathcal{F},$$

quindi si ha la tesi.

(3 \Rightarrow 2) Ovvio.

(2 \Rightarrow 1) Siano $A, B \in \mathcal{B}$, allora (usando che $m(X) = 1$):

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} m(T^{-i}A \cap B) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} (m \times m)\left((T \times T)^{-i}(A \times X) \cap (B \times X)\right) = \\ &= (m \times m)(A \times X)(m \times m)(B \times X) = m(A)m(B), \end{aligned}$$

inoltre

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(m(T^{-i}A \cap B)\right)^2 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} (m \times m)\left((T \times T)^{-i}(A \times A) \cap (B \times B)\right) = \\ &= (m \times m)(A \times A)(m \times m)(B \times B) = \\ &= m(A)^2 m(B)^2. \end{aligned}$$

Dunque

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(m(T^{-i}A \cap B) - m(A)m(B)\right)^2 &= \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(m(T^{-i}A \cap B)^2 - 2m(T^{-i}A \cap B)m(A)m(B) + m(A)^2 m(B)^2\right) = \\ &= 2m(A)^2 m(B)^2 - 2m(A)^2 m(B)^2 = 0, \end{aligned}$$

da cui segue la tesi per il corollario 1.5.6.

□

Capitolo 2

Isomorfismi e spettro di trasformazioni

2.1 Spettro continuo e proprietà ergodiche

Definizione 2.1.1. Si dice che una trasformazione $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ ha lo **spettro continuo** se 1 è il solo autovalore di T (cioè dell'operatore associato definito su $L^2(X)$) e le sole autofunzioni sono le costanti.

Osservazione 2.1.2.

1. T ha lo spettro continuo se e solo se 1 è il solo autovalore e T è ergodico.
2. Siccome l'operatore U associato a T è unitario, se λ è autovalore di T , allora $|\lambda| = 1$, infatti:

$$\|f\|^2 = \|Uf\|^2 = \langle Uf, Uf \rangle = \langle \lambda f, \lambda f \rangle = |\lambda|^2 \|f\|^2 \quad \forall f \in L^2.$$

3. $\lambda = 1$ è sempre un autovalore corrispondente a una/ogni autofunzione costante.

Teorema 2.1.3 (spettrale per operatori unitari).

Se U è un operatore unitario sullo spazio di Hilbert complesso \mathcal{H} , allora per ogni $f \in \mathcal{H}$ esiste un'unica misura di Borel finita μ_f definita su $S^1 \subset \mathbb{C}$ tale che:

$$\langle U^n f, f \rangle = \int_{S^1} \lambda^n d\mu_f(\lambda) \quad \forall n \in \mathbb{Z}.$$

Corollario 2.1.4. Se $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ è invertibile e preserva la misura, allora l'operatore associato $U_T : L^2(X) \rightarrow L^2(X)$ è unitario e se T ha lo spettro continuo allora per ogni $f \in L^2(X)$ tale che $\langle f, 1 \rangle = 0$ la misura μ_f non è atomica.

Teorema 2.1.5. Se T è una trasformazione che preserva la misura e invertibile su uno spazio di probabilità (X, \mathcal{B}, m) , allora:

$$T \text{ weak-mixing} \iff T \text{ ha lo spettro continuo.}$$

Dimostrazione.

(\Rightarrow) Sia $U(f) = f \circ T = \lambda f$ (q.o.) con $f \in L^2$. Se $|\lambda| = 1$, ma $\lambda \neq 1$, allora integrando si ha

$$\int_X f(x) dm(x) = \int_X f(Tx) dm(x) = \int_X \lambda f(x) dm(x),$$

cioè $\int_X f dm = 0$, cioè $\langle f, 1 \rangle = 0$. Da ciò segue per la proprietà di weak-mixing che:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} |\langle U^i f, f \rangle| = 0, \text{ ovvero } \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} |\langle \lambda^i f, f \rangle| = 0,$$

ma siccome $|\lambda^i| = 1$, allora $\langle f, f \rangle = 0$ e dunque $f = 0$ (q.o.).

Invece nel caso in cui $\lambda = 1$ allora $f = \text{costante}$ (q.o.) per l'ergodicità di T .

(\Leftarrow) Innanzitutto si può osservare che per il teorema 1.5.7 la tesi equivale a mostrare che per ogni $f \in L^2$

$$\frac{1}{n} \sum_{i=0}^{n-1} \left| \langle U^i f, f \rangle - \langle f, 1 \rangle \langle 1, f \rangle \right|^2 \xrightarrow{n \rightarrow \infty} 0. \quad (2.1)$$

Se $f \equiv c$ (q.o.) allora la (2.1) vale in quanto:

$$\frac{1}{n} \sum_{i=0}^{n-1} \left| \langle U^i c, c \rangle - \langle c, 1 \rangle \langle 1, c \rangle \right|^2 = \frac{1}{n} \sum_{i=0}^{n-1} |c^2 - c^2|^2 = 0.$$

A questo punto, basta mostrare che:

$$\langle f, 1 \rangle = 0 \implies \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} |\langle U^i f, f \rangle|^2 = 0,$$

inoltre, per il teorema spettrale 2.1.3 basta mostrare che se μ_f è una misura continua (cioè non atomica) su S^1 allora

$$\frac{1}{n} \sum_{i=0}^{n-1} \left| \int_{S^1} \lambda^i d\mu_f(\lambda) \right|^2 \xrightarrow{n \rightarrow \infty} 0. \quad (2.2)$$

Ma infatti, siccome (usando il teorema di Fubini-Tonelli)

$$\begin{aligned} \frac{1}{n} \sum_{i=0}^{n-1} \left| \int_{S^1} \lambda^i d\mu_f(\lambda) \right|^2 &= \frac{1}{n} \sum_{i=0}^{n-1} \left[\int_{S^1} \lambda^i d\mu_f(\lambda) \int_{S^1} \lambda^{-i} d\mu_f(\lambda) \right] = \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \left[\left(\int_{S^1} \lambda^i d\mu_f(\lambda) \right) \left(\int_{S^1} \tau^{-i} d\mu_f(\tau) \right) \right] \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \iint_{S^1 \times S^1} (\lambda \bar{\tau})^i d(\mu_f \times \mu_f)(\lambda, \tau) = \\ &= \iint_{S^1 \times S^1} \left(\frac{1}{n} \sum_{i=0}^{n-1} (\lambda \bar{\tau})^i \right) d(\mu_f \times \mu_f)(\lambda, \tau) \end{aligned} \quad (2.3)$$

allora, se (λ, τ) non è nella diagonale Δ di $S^1 \times S^1$ (cioè $\lambda \neq \tau$), si ha

$$\frac{1}{n} \sum_{i=0}^{n-1} (\lambda\bar{\tau})^i = \frac{1}{n} \left[\frac{1 - (\lambda\bar{\tau})^n}{1 - (\lambda\bar{\tau})} \right] \xrightarrow{n \rightarrow \infty} 0.$$

Ora, siccome μ_f è una misura non atomica, allora $(\mu_f \times \mu_f)(\Delta) = 0$, quindi l'ultimo integrando della (2.3) tende a 0 q.o. e, siccome tale integrando ha modulo ≤ 1 , allora applicando il teorema della convergenza dominata si ottiene quanto voluto in (2.2). \square

Esempio 2.1.6. L'identità id su uno spazio di misura (X, \mathcal{B}, m) è ergodica se e solo se è strong-mixing se e solo se ogni elemento di \mathcal{B} ha misura 0 oppure 1.

Esempio 2.1.7. Una rotazione $T : S^1 \rightarrow S^1$ data da $T(z) = az$ non è mai weak-mixing perché se $f(z) = z$ allora $f(Tz) = az = af(z)$ e $f \neq \text{costante}$.

Esempio 2.1.8. Sia G un gruppo (abeliano) topologico compatto e metrico e sia $A : G \rightarrow G$ un endomorfismo (cioè omomorfismo e continuo). Allora:

$$A \text{ ergodico} \iff A \text{ weak-mixing} \iff A \text{ strong-mixing.}$$

Dimostrazione. Ovviamente, basta mostrare che se A è ergodico allora è anche strong-mixing.

Siano $\gamma, \delta \in \widehat{G}$ caratteri di G . Siccome (osservazione 1.4.12) i caratteri sono un set ortonormale completo di L^2 allora $\langle U_A^n \gamma, \delta \rangle = 0$, a meno che $\gamma = \delta \equiv 1$. In ogni caso si ha che

$$\langle U_A^n \gamma, \delta \rangle \xrightarrow{n \rightarrow \infty} \langle \gamma, 1 \rangle \langle \delta, 1 \rangle = 0.$$

Fissato $\delta \in \widehat{G}$, l'insieme $\mathcal{H}_\delta = \{f \in L^2 \mid \langle U_A^n \gamma, \delta \rangle \rightarrow_n \langle \gamma, 1 \rangle \langle \delta, 1 \rangle\}$ è un sottospazio chiuso di L^2 , infatti se $\{f_k\}_k \subset \mathcal{H}_\delta$ con $f_k \rightarrow_k f$

- per $\delta \equiv 1$ (o una qualsiasi costante) è chiaro che $\mathcal{H}_\delta = L^2$;
- per $\langle \delta, 1 \rangle = 0$ (cioè δ nel sottospazio ortogonale a quello delle costanti), fissato $\varepsilon > 0$ esiste $k \in \mathbb{N}$ tale che $\|f - f_k\|_2 < \varepsilon/2$ ed esiste $N \in \mathbb{N}$ per cui per ogni $n \geq N$ si ha $|\langle U_A^n f_k, \delta \rangle| < \varepsilon/2$, dunque

$$\begin{aligned} \forall n \geq N \quad |\langle U_A^n f, \delta \rangle| &\leq |\langle U_A^n f, \delta \rangle - \langle U_A^n f_k, \delta \rangle| + |\langle U_A^n f_k, \delta \rangle| \leq \\ &\leq \|f - f_k\|_2 \|\delta\|_2 + |\langle U_A^n f_k, \delta \rangle| = \\ &= \|f - f_k\|_2 + |\langle U_A^n f_k, \delta \rangle| < \varepsilon \end{aligned}$$

Ora, siccome $\widehat{G} \subseteq \mathcal{H}_\delta$, allora anche in questo caso $\mathcal{H}_\delta = L^2$.

Ora, presa $f \in L^2$ e posto $\mathcal{L}_f := \{g \in L^2 \mid \langle U_A^n f, g \rangle \rightarrow_n \langle f, 1 \rangle \langle g, 1 \rangle\}$, allora \mathcal{L}_f è un sottospazio chiuso di L^2 contenente \widehat{G} , quindi $\mathcal{L}_f = L^2$ e dunque:

$$\forall f, g \in L^2 \quad \langle U_A^n f, g \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle g, 1 \rangle,$$

cioè A è strong-mixing. \square

2.2 Trasformazioni isomorfe e coniugate

Alcuni dei problemi in teoria della misura sorgono a causa degli insiemi di misura nulla. Si consideri uno spazio di probabilità (X, \mathcal{B}, m) e si consideri l'insieme quoziente \mathcal{C} costituito dalle classi di equivalenza secondo la relazione:

$$\forall A, B \in \mathcal{B} \quad A \sim B \iff m(A \Delta B) = 0.$$

Si può verificare che \mathcal{C} è un'algebra booleana con le operazioni di intersezione, unione (anche numerabili) e complemento ereditate da \mathcal{B} . Inoltre, siccome da $m(A \Delta B) = 0$ segue $m(A) = m(B)$, allora si può considerare m come definita su \mathcal{C} e la struttura (\mathcal{C}, m) è una σ -algebra booleana.

Ora, presa T trasformazione che preserva la misura su X , essa passa al quoziente se si definisce la mappa \mathcal{C} in sé $[E] \mapsto [T^{-1}E]$ (che sarà indicata con T^{-1}); inoltre anch'essa preserva la misura e le operazioni booleane.

A questo punto, un problema fondamentale diventa quando (e in quale senso) considerare equivalenti due trasformazioni tra spazi di misura.

Definizione 2.2.1. *Siano $(X_1, \mathcal{B}_1, m_1)$ e $(X_2, \mathcal{B}_2, m_2)$ due spazi di probabilità e siano $T_1 : X_1 \rightarrow X_1$ e $T_2 : X_2 \rightarrow X_2$ trasformazioni che preservano la misura. Si dice che T_1 è **isomorfa** a T_2 ($T_1 \simeq T_2$) se esistono $M_1 \in \mathcal{B}_1$ e $M_2 \in \mathcal{B}_2$ tali che:*

- $m(M_1) = 1 = m(M_2)$,
- $T_1 M_1 \subseteq M_1, T_2 M_2 \subseteq M_2$
- esiste $\phi : M_1 \rightarrow M_2$ invertibile e che preserva la misura tale che $\phi \circ T_1 = T_2 \circ \phi$.

Osservazione 2.2.2.

1. La relazione \simeq di isomorfismo è una relazione di equivalenza.
2. $T_1 \simeq T_2 \implies T_1^n \simeq T_2^n \quad \forall n$.
3. Se T_1, T_2 sono invertibili, allora si possono prendere M_1, M_2 in modo che $T_i M_i = M_i$ (per $i = 1, 2$), infatti basta scegliere $M_i = \bigcap_{k \in \mathbb{Z}} T_i^k M_i$.

Definizione 2.2.3. *Sia (X, \mathcal{B}, m) uno spazio di misura.*

1. *Si definisca la relazione d'equivalenza*

$$A \sim B \iff m(A \Delta B) = 0,$$

con $(A, B \in \mathcal{B})$ e si denoti l'insieme quoziente con $\tilde{\mathcal{B}} := \mathcal{B} / \sim$, che è una σ -algebra booleana con le operazioni indotte da \mathcal{B} ; inoltre m induce un'applicazione $\tilde{m} : \tilde{\mathcal{B}} \rightarrow [0, 1]$ definita da $\tilde{m}(\tilde{B}) = m(B)$, dove la classe d'equivalenza è $[B] = \tilde{B} \in \tilde{\mathcal{B}}$.

Questa struttura $(\tilde{\mathcal{B}}, \tilde{m})$ viene definita **algebra di misura**.

2. *Sia $T : X \rightarrow X$ è una trasformazione che preserva la misura. Se $A \sim B$ allora $T^{-1}A \sim T^{-1}B$, quindi è ben definita l'applicazione $\tilde{T}^{-1} : \tilde{\mathcal{B}} \rightarrow \tilde{\mathcal{B}}$ come $\tilde{T}^{-1}(\tilde{B}) = \widetilde{T^{-1}(B)}$ ed è una trasformazione che preserva unioni, intersezioni, complementi ed è tale che $\tilde{m}(\tilde{T}^{-1}(\tilde{B})) = \tilde{m}(\tilde{B})$.*

Definizione 2.2.4. Una mappa $\Phi : (\tilde{\mathcal{B}}_2, \tilde{m}_2) \rightarrow (\tilde{\mathcal{B}}_1, \tilde{m}_1)$ tra algebre di misura è detta **isomorfismo di algebre di misura** se è bigettiva, preserva complemento e unione numerabile ed è tale che $\tilde{m}_1(\Phi(\tilde{B})) = \tilde{m}_2(\tilde{B})$ per ogni $\tilde{B} \in \tilde{\mathcal{B}}_2$.

Dalla definizione 2.2.3 segue che data $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ si può considerare $\tilde{T}^{-1} : (\tilde{\mathcal{B}}, \tilde{m}) \rightarrow (\tilde{\mathcal{B}}, \tilde{m})$. Si noti che se $B \in \mathcal{B}$ è un rappresentante della classe di equivalenza $\tilde{B} \in \tilde{\mathcal{B}}$, allora anche $\chi_{\tilde{B}}$ è un elemento ben definito di $L^2(X, \mathcal{B}, m)$.

Qual è il rapporto tra \tilde{T}^{-1} e l'applicazione indotta definita da $\chi_{\tilde{B}} \mapsto \chi_{\tilde{T}^{-1}(\tilde{B})}$?
E nel caso in cui \tilde{T}^{-1} sia un isomorfismo di algebre di misura?

Proposizione 2.2.5. Un operatore unitario $U : L^2(X, \mathcal{B}, m) \rightarrow L^2(X, \mathcal{B}, m)$ è indotto da un isomorfismo di algebre di misura $\tilde{T}^{-1} : (\tilde{\mathcal{B}}, \tilde{m}) \rightarrow (\tilde{\mathcal{B}}, \tilde{m})$ se e soltanto se U e U^{-1} mandano funzioni limitate in funzioni limitate e se vale $U(f \cdot g) = U(f) \cdot U(g)$ per f, g funzioni limitate.

Dimostrazione.

(\Rightarrow) Se U è indotto dall'isomorfismo \tilde{T}^{-1} (cioè dalla trasformazione invertibile e che preserva la misura $T : X \rightarrow X$) allora vale che:

$$f \text{ limitata} \implies Uf = f \circ T \text{ limitata}$$

$$U(f \cdot g) = (f \cdot g) \circ T = (f \circ T) \cdot (g \circ T) = Uf \cdot Ug.$$

(\Leftarrow) Se $f = \chi_B$ con $B \in \mathcal{B}$ è una funzione indicatrice, allora $f^2 = f$ e quindi $(Uf)^2 = (f \circ T)^2 = f \circ T = Uf$, quindi anche Uf è una funzione indicatrice (proprio perché assume come valori solo 0 o 1) corrispondente ad un opportuno $G \in \mathcal{B}$.

Prendendo $TB = G$ la mappa $T : B \mapsto TB = G$ è definita da $\tilde{\mathcal{B}}$ in sé; inoltre T è surgettiva, in quanto U lo è, e siccome U è iniettiva, cioè $\text{Ker}U = \{0\}$, allora vale che $TB = 0 \implies B = 0$ (dove 0 indica l'elemento nullo in $\tilde{\mathcal{B}}$).

Resta da verificare che T è una trasformazione che preserva la misura ed è σ -omomorfismo. Ma infatti:

- U unitario $\implies U$ preserva la norma $\implies T$ preserva la misura poiché $m(B) = \|\chi_B\|^2$.
- T preserva le intersezioni perché U è moltiplicativo e vale $\chi_{A \cap B} = \chi_A \cdot \chi_B$. Inoltre, T preserva le unioni perché vale che $\chi_{A \cup B} = \chi_A + \chi_B - \chi_A \cdot \chi_B$ e quindi usando un argomento induttivo e la continuità di U si dimostra che T preserva le unioni numerabili.

□

Definizione 2.2.6. Si dice che $T_1 : (X_1, \mathcal{B}_1, m_1) \rightarrow (X_1, \mathcal{B}_1, m_1)$ e $T_2 : (X_2, \mathcal{B}_2, m_2) \rightarrow (X_2, \mathcal{B}_2, m_2)$ sono **trasformazioni coniugate** se esiste un isomorfismo di algebre di misure $\Phi : (\tilde{\mathcal{B}}_2, \tilde{m}_2) \rightarrow (\tilde{\mathcal{B}}_1, \tilde{m}_1)$ tale che $\Phi \circ \tilde{T}_2^{-1} = \tilde{T}_1^{-1} \circ \Phi$.

Osservazione 2.2.7.

1. Il coniugio è una relazione d'equivalenza.
2. $T_1 \simeq T_2$ isomorfe $\implies T_1, T_2$ coniugate.
Infatti, siccome esiste ϕ tale che $\phi \circ T_1 = T_2 \circ \phi$, allora basta porre $\Phi = \tilde{\phi}^{-1}$ (e si noti che è ben definita anche se ϕ non è definita su tutto X_1).

2.3 Isomorfismi spettrali e invarianti spettrali

Definizione 2.3.1. Si dice che $T_1 : (X_1, \mathcal{B}_1, m_1) \rightarrow (X_1, \mathcal{B}_1, m_1)$ e $T_2 : (X_2, \mathcal{B}_2, m_2) \rightarrow (X_2, \mathcal{B}_2, m_2)$ sono **spettralmente isomorfe** se esiste un operatore lineare $W : L^2(X_1, \mathcal{B}_1, m_1) \rightarrow L^2(X_2, \mathcal{B}_2, m_2)$ tale che:

1. W è invertibile;
2. $\langle Wf, Wg \rangle_2 = \langle f, g \rangle_1$ per ogni $f, g \in L^2(X_1, \mathcal{B}_1, m_1)$;
3. $U_1 \circ W = W \circ U_2$, dove U_i è l'operatore associato a T_i .

(Le sole condizioni (1) e (2) definiscono un **isomorfismo di spazi di Hilbert**).

Osservazione 2.3.2.

1. L'isomorfismo spettrale è una relazione d'equivalenza.
2. Se $\Phi : (\tilde{\mathcal{B}}_2, \tilde{m}_2) \rightarrow (\tilde{\mathcal{B}}_1, \tilde{m}_1)$ è un isomorfismo di algebre di misura, allora esso induce una mappa invertibile lineare $V : L^2(X_2, \mathcal{B}_2, m_2) \rightarrow L^2(X_1, \mathcal{B}_1, m_1)$ con le proprietà:
 - $\langle Vf, Vg \rangle_1 = \langle f, g \rangle_2$ per ogni $f, g \in L^2(X_2, \mathcal{B}_2, m_2)$
 - V, V^{-1} mappano funzioni limitate in funzioni limitate
 - V è moltiplicativo sulle funzioni limitate, cioè: $V(fg) = V(f)V(g)$ per ogni f, g funzioni limitate.

Dimostrazione. Definita da $V(\chi_{\tilde{B}}) = \chi_{\Phi(\tilde{B})}$ sulle funzioni indicatrici, si estende per linearità e continuità prima alle funzioni semplici e poi a tutto $L^2(X_2, \mathcal{B}_2, m_2)$. (Si veda anche la proposizione 2.2.5). \square

3. T_1, T_2 coniugate $\implies T_1, T_2$ spettralmente isomorfe.

Dimostrazione. Per ipotesi esiste $\Phi : (\tilde{\mathcal{B}}_2, \tilde{m}_2) \rightarrow (\tilde{\mathcal{B}}_1, \tilde{m}_1)$, isomorfismo di algebre di misura, tale che $\Phi \circ \tilde{T}_2^{-1} = \tilde{T}_1^{-1} \circ \Phi$. Sia V definita come al punto precedente $V(\chi_{\tilde{B}}) = \chi_{\Phi(\tilde{B})}$.

Resta da verificare che $V \circ U_2 = U_1 \circ V$. Innanzitutto, ciò vale per funzioni indicatrici

$$\begin{aligned} U_1 \circ V(\chi_B) &= U_1(\chi_{\Phi(B)}) = \chi_{\Phi(B)} \circ T_1 = \chi_{T_1^{-1}(\Phi(B))} = \chi_{\tilde{T}_1^{-1}(\Phi(B))} = \\ &= \chi_{\Phi(\tilde{T}_2^{-1}(B))} = \chi_{\Phi(T_2^{-1}(B))} = V(\chi_{T_2^{-1}(B)}) = V(\chi_B \circ T_2) = \\ &= V \circ U_2(\chi_B), \end{aligned}$$

poi per linearità e continuità si estende alle funzioni semplici e a tutto L^2 . \square

Proposizione 2.3.3. *Un isomorfismo (di spazi di Hilbert) $V : L^2(X_2, \mathcal{B}_2, m_2) \rightarrow L^2(X_1, \mathcal{B}_1, m_1)$ è indotto da un isomorfismo di algebre di misura se entrambi V, V^{-1} mandano funzioni limitate in funzioni limitate e se V è moltiplicativo sulle funzioni limitate.*

Dimostrazione. Si veda la dimostrazione della proposizione 2.2.5. \square

Corollario 2.3.4. *Se $T_1 : (X_1, \mathcal{B}_1, m_1) \rightarrow (X_1, \mathcal{B}_1, m_1)$ e $T_2 : (X_2, \mathcal{B}_2, m_2) \rightarrow (X_2, \mathcal{B}_2, m_2)$ sono trasformazioni che preservano la misura e se $U_1 \circ V = V \circ U_2$ con $V : L^2(X_2, \mathcal{B}_2, m_2) \rightarrow L^2(X_1, \mathcal{B}_1, m_1)$ isomorfismo di spazi di Hilbert, allora T_1 e T_2 sono coniugate.*

Una volta definite le classi di equivalenza tra trasformazioni, è naturale chiedersi se vi siano delle proprietà (e/o quantità) che sono invarianti all'interno di queste classi; ovvero, si cercano delle proprietà P tali che se T_1 gode della proprietà P e se T_2 è nella stessa classe di equivalenza di T_1 , allora anche T_2 gode della P .

Teorema 2.3.5. *Le seguenti proprietà sono invarianti spettrali (e quindi anche invarianti per coniugio e isomorfismo):*

1. ergodicità,
2. weak-mixing,
3. strong-mixing.

Dimostrazione.

1. T ergodica $\iff \{f \in L^2 \mid Uf = f\}$ è un sottospazio 1-dimensionale.
2. T weak-mixing $\iff 1$ è l'unico autovalore e T è ergodica.
3. Sia T_1 trasformazione strong-mixing e sia T_2 una trasformazione spettralmente isomorfa a T_1 . Quindi per ipotesi esiste W tale che $W \circ U_2 = U_1 \circ W$. La tesi equivale a mostrare che

$$\langle U_2^n f, g \rangle \xrightarrow{n \rightarrow \infty} \langle f, 1 \rangle \langle 1, g \rangle \quad \forall f, g \in L^2(X_2, \mathcal{B}_2, m_2). \quad (2.4)$$

La (2.4) è valida se f (oppure g) è costante.

Quindi si può assumere che $\langle f, 1 \rangle = 0 = \langle g, 1 \rangle$, così che la 2.4 diventa:

$$\langle U_2^n f, g \rangle \xrightarrow{n \rightarrow \infty} 0 \quad \forall f, g \in L^2(X_2, \mathcal{B}_2, m_2).$$

Siccome T_1 è ergodica, allora anche T_2 è ergodica (per il punto (1) appena dimostrato); inoltre, siccome W manda le funzioni T_2 -invarianti surgettivamente in quelle T_1 -invarianti allora W manda le costanti di $L^2(X_2, \mathcal{B}_2, m_2)$ nelle costanti di $L^2(X_1, \mathcal{B}_1, m_1)$ surgettivamente. Da ciò segue che $\langle Wf, 1 \rangle = 0 = \langle Wg, 1 \rangle$.

Infine, siccome W preserva il prodotto scalare, T_1 è strong-mixing e $W \circ U_2 = U_1 \circ W$, allora si ha:

$$\langle U_2^n f, g \rangle = \langle WU_2^n f, Wg \rangle = \langle U_1^n Wf, Wg \rangle \xrightarrow{n \rightarrow \infty} \langle Wf, 1 \rangle \langle 1, Wg \rangle = 0.$$

Dunque T_2 è strong-mixing.

□

Esempio 2.3.6. Siano $T_1, T_2 : S^1 \rightarrow S^1$ rotazioni della forma $T_1(z) = a_1 z$ e $T_2(z) = a_2 z$, con a_1 che è radice dell'unità ma a_2 che non lo è. Allora, T_1 non è ergodica mentre T_2 lo è, quindi non possono essere spettralmente isomorfe.

Definizione 2.3.7. Una trasformazione invertibile e che preserva la misura $T : X \rightarrow X$ ha uno **spettro numerabile di Lebesgue** se esiste una base ortonormale di $L^2(X)$ della forma: $f_0 \equiv 1$, $\{U^n f_i\}_{i \in \mathbb{N}^+, n \in \mathbb{Z}}$. Ovvero:

$$\begin{aligned} f_0 &\equiv 1 \\ \dots, U^{-1}f_1, f_1, Uf_1, U^2f_1, \dots \\ \dots, U^{-1}f_2, f_2, Uf_2, U^2f_2, \dots \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \end{aligned}$$

Proposizione 2.3.8. Due trasformazioni con lo spettro numerabile di Lebesgue sono sempre spettralmente isomorfe.

Dimostrazione. Siano $T : X \rightarrow X$ e $S : Y \rightarrow Y$ le due trasformazioni, con basi ortonormali rispettivamente date da

$$f_0 \equiv 1, \{U_T^n f_i\}_{i \in \mathbb{N}^+, n \in \mathbb{Z}} \quad \text{e} \quad g_0 \equiv 1, \{U_S^n g_i\}_{i \in \mathbb{N}^+, n \in \mathbb{Z}}.$$

A questo punto si definisca $W : L^2(Y) \rightarrow L^2(X)$ ponendo

$$g_0 \mapsto f_0 \quad , \quad U_S^n g_i \mapsto U_T^n f_i$$

ed estesa per linearità a tutto $L^2(Y)$.

Allora, $W \circ U_S = U_T \circ W$ e S, T sono spettralmente isomorfe. □

Teorema 2.3.9. Se T ha lo spettro numerabile di Lebesgue allora è strong-mixing.

Dimostrazione. Per ipotesi esiste una base ortonormale $f_0 \equiv 1$, $\{U^n f_i\}_{i \in \mathbb{N}^+, n \in \mathbb{Z}}$. Allora

$$\langle (U^p \circ U^n) f_m, U^k f_q \rangle \xrightarrow{p \rightarrow \infty} \langle U^n f_m, 1 \rangle \langle 1, U^k f_q \rangle \quad \forall k, n \in \mathbb{Z}, \forall m, q \in \mathbb{N}.$$

Fissati k, q si definisca $\mathcal{H}_{k,q} := \{f \in L^2 \mid \langle U^p f, U^k f_q \rangle \rightarrow_p \langle f, 1 \rangle \langle 1, U^k f_q \rangle\}$, che è un sottospazio chiuso di L^2 e contenente la base $\{f_0, \{U^n f_i\}_{i \in \mathbb{N}^+, n \in \mathbb{Z}}\}$, quindi $\mathcal{H}_{k,q} = L^2$.

Ora, fissata $f \in L^2$ e posto $\mathcal{L}_f := \{g \in L^2 \mid \langle U^p f, g \rangle \rightarrow_p \langle f, 1 \rangle \langle 1, g \rangle\}$, anch'esso è un sottospazio chiuso di L^2 contenente la base ortonormale come sopra, quindi $\mathcal{L}_f = L^2$. Dunque

$$\langle U^p f, g \rangle \xrightarrow{p \rightarrow \infty} \langle f, 1 \rangle \langle 1, g \rangle \quad \forall f, g \in L^2,$$

cioè T è strong-mixing. □

Esempio 2.3.10. Si considerino il (p_1, \dots, p_n) -shift (bilatero) e il (q_1, \dots, q_m) -shift (bilatero) nell'ipotesi in cui almeno due tra i numeri p_i siano non nulli e lo stesso valga per q_j . Allora questi shifts sono spettralmente isomorfi ma non necessariamente coniugati (con l'introduzione del concetto di entropia vedremo che essi non possono essere coniugati).

Si consideri, ad esempio, il caso particolare dello $(\frac{1}{2}, \frac{1}{2})$ -shift T sullo spazio $X = \prod_{-\infty}^{+\infty} \{-1, 1\}$. Una base di $L^2(\{-1, 1\})$ è data dalla funzione costante 1 e dall'identità id . Inoltre, $L^2(X)$ è il prodotto degli spazi $L^2(\{-1, 1\})$ e quindi una base ortonormale per $L^2(X)$ è costituita dalle funzioni da X in \mathbb{C} della forma:

$$\begin{aligned} g_0(\{x_n\}) &= 1 \\ g_{n_1, \dots, n_r}(\{x_n\}) &= x_{n_1} \cdot x_{n_2} \cdot \dots \cdot x_{n_r} \quad , \quad \forall n_1 < \dots < n_r . \end{aligned}$$

Si noti che (detto U l'operatore associato a T)

$$\begin{aligned} U(g_{n_1, \dots, n_r})(\{x_n\}) &= (g_{n_1, \dots, n_r} \circ T)(\{x_n\}) = \\ &= x_{n_1+1} \cdot \dots \cdot x_{n_r+1} = g_{n_1+1, \dots, n_r+1}(\{x_n\}) \quad , \end{aligned}$$

cioè

$$U(g_{n_1, \dots, n_r}) = g_{n_1+1, \dots, n_r+1} .$$

Dunque T ha lo spettro numerabile di Lebesgue e quindi è strong-mixing (e analogamente si può verificare per ogni generico (p_1, \dots, p_n) -shift). Inoltre, sempre dai risultati precedenti segue che i due shifts considerati sopra sono spettralmente isomorfi.

Capitolo 3

Entropia

3.1 Nozioni preliminari

Definizione 3.1.1. Una **partizione** di uno spazio di probabilità (X, \mathcal{B}, m) è una collezione di elementi disgiunti di \mathcal{B} la cui unione dà X .

Se ξ è una partizione finita di (X, \mathcal{B}, m) allora la collezione degli insiemi formati da unioni di elementi di ξ è detta **sotto- σ -algebra** ed è indicata con $\mathcal{A}(\xi)$.

Viceversa, se $\mathcal{C} = \{C_i \mid i = 1, \dots, n\}$ è una sotto- σ -algebra (finita) di \mathcal{B} , allora gli insiemi (non vuoti) della forma $B_1 \cap \dots \cap B_n$, dove $B_i = C_i$ oppure $B_i = C_i^c$, sono una partizione finita di X che è indicata con $\xi(\mathcal{C})$.

Dunque vi è una corrispondenza binivoca tra partizioni finite e sotto- σ -algebre (finite).

Definizione 3.1.2. Se ξ, η sono partizioni finite, allora si dice che η è un **raffinamento** di ξ , indicato con $\xi \leq \eta$, se ogni elemento di ξ è unione di elementi di η . Si noti che:

$$\xi \leq \eta \iff \mathcal{A}(\xi) \subseteq \mathcal{A}(\eta)$$

$$\mathcal{A} \subseteq \mathcal{C} \iff \xi(\mathcal{A}) \leq \xi(\mathcal{C}).$$

Osservazione 3.1.3.

1. Siano $\xi = \{A_1, \dots, A_n\}$, $\eta = \{C_1, \dots, C_k\}$ partizioni di X , e si indichi con $\xi \vee \eta := \{A_i \cap C_j \mid 1 \leq i \leq n, 1 \leq j \leq k\}$.

Siano \mathcal{A}, \mathcal{C} sotto- σ -algebre finite di \mathcal{B} e si indichi con $\mathcal{A} \vee \mathcal{C}$ la più piccola sotto- σ -algebra di \mathcal{B} contenente \mathcal{A} e \mathcal{C} .

Si noti che

$$\xi(\mathcal{A} \vee \mathcal{C}) = \xi(\mathcal{A}) \vee \xi(\mathcal{C})$$

$$\mathcal{A}(\xi \vee \eta) = \mathcal{A}(\xi) \vee \mathcal{A}(\eta).$$

2. Siano $T : X \rightarrow X$ una trasformazione che preserva la misura, $\xi = \{A_1, \dots, A_m\}$ una partizione di X e si definiscano $T^{-n}\xi := \{T^{-n}A_1, \dots, T^{-n}A_m\}$ e $T^{-n}(\mathcal{A}) := \{T^{-n}A \mid A \in \mathcal{A}\}$ (con $n \in \mathbb{N}$). Si noti che:

- $\xi(T^{-n}(\mathcal{A})) = T^{-n}\xi(\mathcal{A})$
- $\mathcal{A}(T^{-n}\xi) = T^{-n}(\mathcal{A}(\xi))$

- $T^{-n}(\mathcal{A} \vee \mathcal{C}) = T^{-n}(\mathcal{A}) \vee T^{-n}(\mathcal{C})$
- $T^{-n}(\xi \vee \eta) = T^{-n}(\xi) \vee T^{-n}(\eta)$
- $\xi \leq \eta \implies T^{-n}(\xi) \leq T^{-n}(\eta)$
- $\mathcal{A} \subseteq \mathcal{C} \implies T^{-n}(\mathcal{A}) \subseteq T^{-n}(\mathcal{C})$.

Definizione 3.1.4. Se \mathcal{D}, \mathcal{E} sono sotto- σ -algebre (non necessariamente finite) di X allora si definiscono le relazioni:

$$\mathcal{D} \dot{\subset} \mathcal{E} \iff \forall D \in \mathcal{D}, \exists E \in \mathcal{E} : m(D \triangle E) = 0$$

$$\mathcal{D} \dot{=} \mathcal{E} \iff \mathcal{D} \dot{\subset} \mathcal{E} \text{ e } \mathcal{E} \dot{\subset} \mathcal{D}.$$

Si noti che se \mathcal{D}, \mathcal{E} sono finite e $\mathcal{D} \dot{\subset} \mathcal{E}$ con $\xi(\mathcal{D}) = \{D_1, \dots, D_p, D_{p+1}, \dots, D_q\}$, dove $m(D_i) > 0$ per $1 \leq i \leq p$ e $m(D_i) = 0$ per $p+1 \leq i \leq q$, allora

$$\xi(\mathcal{E}) = \{E_1, \dots, E_p, E_{p+1}, \dots, E_s\}$$

con $m(D_i \triangle E_i) = 0$ per $1 \leq i \leq p$ e $m(E_i) = 0$ per $p+1 \leq i \leq s$.

3.2 Entropia

Definizione 3.2.1. Sia \mathcal{A} una sotto- σ -algebra finita di \mathcal{B} e sia $\xi(\mathcal{A}) = \{A_1, \dots, A_k\}$ la partizione associata. Si definisce l'**entropia di \mathcal{A} (o di $\xi(\mathcal{A})$)** come

$$H(\mathcal{A}) = H(\xi(\mathcal{A})) = - \sum_{i=1}^k m(A_i) \log_2 m(A_i).$$

(In seguito si adotterà la convenzione che $0 \cdot \log_2 0 = 0$)

Osservazione 3.2.2.

1. Se $\mathcal{A} = \{\emptyset, X\}$ allora $H(\mathcal{A}) = 0$.
2. Se $\xi(\mathcal{A}) = \{A_1, \dots, A_k\}$ con $m(A_i) = 1/k \forall i$, allora

$$H(\mathcal{A}) = - \sum_{i=1}^k \frac{1}{k} \log_2 \frac{1}{k} = \log_2 k.$$

3. $H(\mathcal{A}) \geq 0 \quad \forall \mathcal{A}$.
4. Se $\mathcal{A} \dot{=} \mathcal{C}$ allora $H(\mathcal{A}) = H(\mathcal{C})$.

Definizione 3.2.3. Sia $T : X \rightarrow X$ trasformazione che preserva la misura e sia \mathcal{A} sotto- σ -algebra finita di \mathcal{B} . Si definisce l'**entropia di T rispetto a \mathcal{A}** come:

$$h(T, \mathcal{A}) := \lim_{n \rightarrow \infty} \left[\frac{1}{n} H(\mathcal{A} \vee T^{-1}\mathcal{A} \vee \dots \vee T^{-(n-1)}\mathcal{A}) \right] = \lim_{n \rightarrow \infty} \left[\frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{A} \right) \right].$$

Osservazione 3.2.4.

1. $h(T, \mathcal{A}) \geq 0$

2. Gli elementi di $\xi(\bigvee_i T^{-i}\mathcal{A}) = \bigvee_i \xi(T^{-i}\mathcal{A})$ sono tutti gli insiemi della forma $\bigcap_j T^{-1}A_j$, dove $\xi(\mathcal{A}) = \{A_1, \dots, A_k\}$

Definizione 3.2.5. Si definisce l'entropia di T come

$$h(T) = \sup \left\{ h(T, \mathcal{A}) \mid \mathcal{A} \subset \mathcal{B} \text{ sotto-}\sigma\text{-algebra} \right\} .$$

Osservazione 3.2.6.

1. $h(T) \geq 0$ e può assumere valore $+\infty$
2. $h(id_X) = 0$, infatti per ogni \mathcal{A} sotto- σ -algebra finita si ha

$$\begin{aligned} h(id_X, \mathcal{A}) &= \lim_{n \rightarrow \infty} \frac{1}{n} H\left(\mathcal{A} \vee id_X^{-1}\mathcal{A} \vee \dots \vee id_X^{-(n-1)}\mathcal{A}\right) = \\ &= \lim_n \frac{1}{n} H(\mathcal{A} \vee \dots \vee \mathcal{A}) = \lim_n \frac{1}{n} H(\mathcal{A}) = 0 . \end{aligned}$$

Teorema 3.2.7. L'entropia è invariante per coniugio (e quindi per isomorfismo).

Dimostrazione. Siano $T_i : X_i \rightarrow X_i$ con $i = 1, 2$ due trasformazioni che preservano la misura e sia $\Phi : (\mathcal{B}_2, \tilde{m}_2) \rightarrow (\mathcal{B}_1, \tilde{m}_1)$ un isomorfismo di algebre di misura tale che $\Phi \circ \tilde{T}_2^{-1} = \tilde{T}_1^{-1} \circ \Phi$. Sia $\mathcal{A}_2 \subset \mathcal{B}_2$ sotto- σ -algebra finita con $\xi(\mathcal{A}_2) = \{A_1, \dots, A_r\}$ e si scelgano $B_i \in \mathcal{B}_1$ tali che $\tilde{B}_i = \Phi(\tilde{A}_i)$ in modo che $\eta := \{B_1, \dots, B_r\}$ sia una partizione di $(X_1, \mathcal{B}_1, m_1)$; si ponga $\mathcal{A}_1 := \mathcal{A}(\eta)$. Siccome per ogni $q_i \in \{1, \dots, r\}$

$$\Phi\left(\bigcap_{i=0}^{n-1} T_2^{-i} A_{q_i}\right) = \Phi\left(\bigcap_{i=0}^{n-1} \tilde{T}_2^{-i} \tilde{A}_{q_i}\right) = \bigcap_{i=0}^{n-1} \tilde{T}_1^{-i} \Phi(\tilde{A}_{q_i}) = \bigcap_{i=0}^{n-1} \widetilde{T_1^{-i} B_{q_i}} ,$$

allora

$$m\left(\bigcap_{i=0}^{n-1} T_1^{-i} B_{q_i}\right) = m\left(\bigcap_{i=0}^{n-1} T_2^{-i} A_{q_i}\right) .$$

Quindi

$$H\left(\bigvee_{i=0}^{n-1} T_1^{-i} \mathcal{A}_1\right) = H\left(\bigvee_{i=0}^{n-1} T_2^{-i} \mathcal{A}_2\right) ,$$

da cui segue che $h(T_1, \mathcal{A}_1) = h(T_2, \mathcal{A}_2)$, il che implica $h(T_1) \geq h(T_2)$.

Con un ragionamento simmetrico, si deduce la disuguaglianza opposta e quindi si ottiene la tesi $h(T_1) = h(T_2)$. \square

Corollario 3.2.8. Se $\xi = \{A_1, \dots, A_k\}$ è una partizione, allora $H(\xi) \leq \log_2 k$.

Dimostrazione. Presi $a_i = 1/k$ e $x_i = m(A_i)$ con $i = 1, \dots, k$ si ottiene la tesi sfruttando la convessità della funzione $\phi : [0, +\infty) \rightarrow \mathbb{R}$ definita da

$$\phi(x) = \begin{cases} 0 , & \text{se } x = 0 \\ x \log_2 x , & \text{se } x > 0 . \end{cases}$$

\square

In accordo col concetto “fisico-statistico” dell’entropia, dal corollario precedente e dall’osservazione 3.2.2 si ottiene che l’entropia è massima nel caso in cui si abbia una partizione con insiemi di eguale misura. Ovvero, in una situazione in cui le diverse “configurazioni possibili” sono tutte equiprobabili e quindi, in definitiva, c’è difficoltà nel cercare di stabilire delle configurazioni più probabili di un sistema.

3.3 Entropia condizionale

Definizione 3.3.1. Siano $\mathcal{A}, \mathcal{C} \subseteq \mathcal{B}$ sotto- σ -algebre finite con $\xi(\mathcal{A}) = \{A_1, \dots, A_k\}$ e $\xi(\mathcal{C}) = \{C_1, \dots, C_p\}$. Si definisce l’**entropia di \mathcal{A} data \mathcal{C}** come:

$$\begin{aligned} H(\mathcal{A}|\mathcal{C}) &= - \sum_{j=1}^p \left[m(C_j) \sum_{i=1}^k \left[\frac{m(A_i \cap C_j)}{m(C_j)} \log_2 \frac{m(A_i \cap C_j)}{m(C_j)} \right] \right] = \\ &= - \sum_{i,j} \left[m(A_i \cap C_j) \log_2 \frac{m(A_i \cap C_j)}{m(C_j)} \right], \end{aligned}$$

dove si intendono omissi i termini per cui $m(C_j) = 0$.

Dalla definizione di entropia condizionale segue che per calcolare $H(\mathcal{A}|\mathcal{C})$ si considera ogni C_j come spazio di misura con la misura normalizzata $m(\cdot)/m(C_j)$ e poi si calcola l’entropia della partizione di C_j indotta da $\xi(\mathcal{A})$ ed infine si pondera il tutto con $m(C_j)$.

Osservazione 3.3.2.

1. Se $\mathcal{N} = \{\emptyset, X\}$ allora $H(\mathcal{A}|\mathcal{N}) = H(\mathcal{A})$.
2. $H(\mathcal{A}|\mathcal{C}) \geq 0$.
3. Se $\mathcal{A} \doteq \mathcal{D}$ allora $H(\mathcal{A}|\mathcal{C}) = H(\mathcal{D}|\mathcal{C})$.
4. Se $\mathcal{C} \doteq \mathcal{D}$ allora $H(\mathcal{A}|\mathcal{C}) = H(\mathcal{A}|\mathcal{D})$.

Proposizione 3.3.3. Siano $\mathcal{A}, \mathcal{C}, \mathcal{D} \subseteq \mathcal{B}$ sotto- σ -algebre finite. Allora:

1. $H(\mathcal{A} \vee \mathcal{C}|\mathcal{D}) = H(\mathcal{A}|\mathcal{D}) + H(\mathcal{C}|\mathcal{A} \vee \mathcal{D})$.
2. $H(\mathcal{A} \vee \mathcal{C}) = H(\mathcal{A}) + H(\mathcal{C}|\mathcal{A})$.
3. $\mathcal{A} \subseteq \mathcal{C} \implies H(\mathcal{A}|\mathcal{D}) \leq H(\mathcal{C}|\mathcal{D})$.
4. $\mathcal{A} \subseteq \mathcal{C} \implies H(\mathcal{A}) \leq H(\mathcal{C})$.
5. $\mathcal{C} \subseteq \mathcal{D} \implies H(\mathcal{A}|\mathcal{C}) \geq H(\mathcal{A}|\mathcal{D})$.
6. $H(\mathcal{A}) \geq H(\mathcal{A}|\mathcal{D})$.
7. $H(\mathcal{A} \vee \mathcal{C}|\mathcal{D}) \leq H(\mathcal{A}|\mathcal{D}) + H(\mathcal{C}|\mathcal{D})$.
8. $H(\mathcal{A} \vee \mathcal{C}) \leq H(\mathcal{A}) + H(\mathcal{C})$.
9. Se T preserva la misura, allora:

$$H(T^{-1}\mathcal{A}|T^{-1}\mathcal{C}) = H(\mathcal{A}|\mathcal{C}) \quad , \quad H(T^{-1}\mathcal{A}) = H(\mathcal{A}).$$

Dimostrazione. Siano $\xi(\mathcal{A}) = \{A_i\}$, $\xi(\mathcal{C}) = \{C_j\}$ e $\xi(\mathcal{D}) = \{D_k\}$. Si può supporre, senza perdita di generalità, che tutti questi insiemi abbiano misura positiva (poiché se fosse, ad esempio, $m(A_i) = 0$ per $r+1 \leq i \leq k$ allora si potrebbe sostituire $\xi(\mathcal{A})$ con $\{A_1, \dots, A_{r-1}, A_r \cup \dots \cup A_k\}$).

1. Innanzitutto

$$H(\mathcal{A} \vee \mathcal{C} | \mathcal{D}) = - \sum_{i,j,k} m(A_i \cap C_j \cap D_k) \log_2 \frac{m(A_i \cap C_j \cap D_k)}{m(D_k)},$$

ma siccome

$$\frac{m(A_i \cap C_j \cap D_k)}{m(D_k)} = \frac{m(A_i \cap C_j \cap D_k)}{m(A_i \cap D_k)} \frac{m(A_i \cap D_k)}{m(D_k)}$$

eccetto quando $m(A_i \cap D_k) = 0$, ma in tal caso il membro sinistro è nullo e non è considerato in $H(\mathcal{A} \vee \mathcal{C} | \mathcal{D})$, allora da ciò segue la tesi:

$$\begin{aligned} H(\mathcal{A} \vee \mathcal{C} | \mathcal{D}) &= - \sum_{i,j,k} m(A_i \cap C_j \cap D_k) \log \frac{m(A_i \cap D_k)}{m(D_k)} - \\ &\quad - \sum_{i,j,k} m(A_i \cap C_j \cap D_k) \log \frac{m(A_i \cap C_j \cap D_k)}{m(A_i \cap D_k)} = \\ &= - \sum_{i,j,k} m(A_i \cap D_k) \log \frac{m(A_i \cap D_k)}{m(D_k)} + H(\mathcal{C} | \mathcal{A} \vee \mathcal{D}) = \\ &= H(\mathcal{A} | \mathcal{D}) + H(\mathcal{C} | \mathcal{A} \vee \mathcal{D}). \end{aligned}$$

2. Segue dalla (1) prendendo $\mathcal{D} = \{\emptyset, X\}$ e dall'osservazione precedente.

3. Segue dalla (1): $H(\mathcal{C} | \mathcal{D}) = H(\mathcal{A} \vee \mathcal{C} | \mathcal{D}) = H(\mathcal{A} | \mathcal{D}) + H(\mathcal{C} | \mathcal{A} \vee \mathcal{D}) \geq H(\mathcal{A} | \mathcal{D})$.

4. Segue dalla precedente, prendendo $\mathcal{D} = \{\emptyset, X\}$.

5. Fissati i, j e posto $a_k := \frac{m(D_k \cap C_j)}{m(C_j)}$, $x_k := \frac{m(A_i \cap D_k)}{m(D_k)}$, per la convessità della funzione definita su $[0, +\infty)$

$$\phi(x) := x \log x \cdot \chi_{(0, +\infty)}(x)$$

si ha:

$$\phi \left(\sum_k \frac{m(D_k \cap C_j)}{m(C_j)} \frac{m(A_i \cap D_k)}{m(D_k)} \right) \leq \sum_k \frac{m(D_k \cap C_j)}{m(C_j)} \phi \left(\frac{m(A_i \cap D_k)}{m(D_k)} \right),$$

ma siccome per ipotesi $\mathcal{C} \subseteq \mathcal{D}$, allora il primo membro è uguale a

$$\phi \left(\frac{m(A_i \cap C_j)}{m(C_j)} \right) = \frac{m(A_i \cap C_j)}{m(C_j)} \log \frac{m(A_i \cap C_j)}{m(C_j)},$$

allora moltiplicando per $m(C_j)$ e sommando sugli i, j si ottiene:

$$\begin{aligned} \sum_{i,j} m(A_i \cap C_j) \log \frac{m(A_i \cap C_j)}{m(C_j)} &\leq \\ &\leq \sum_{i,j,k} m(D_k \cap C_j) \frac{m(A_i \cap D_k)}{m(D_k)} \log \frac{m(A_i \cap D_k)}{m(D_k)} = \\ &= \sum_{i,k} m(D_k) \frac{m(A_i \cap D_k)}{m(D_k)} \log \frac{m(A_i \cap D_k)}{m(D_k)}, \end{aligned}$$

ovvero $-H(\mathcal{A}|\mathcal{C}) \leq -H(\mathcal{A}|\mathcal{D})$, quindi $H(\mathcal{A}|\mathcal{D}) \leq H(\mathcal{A}|\mathcal{C})$.

6. Segue dalla precedente prendendo $\mathcal{C} = \{\emptyset, X\}$.

7. Segue dalla (1) e dalla (5).

8. Segue dalla precedente prendendo $\mathcal{D} = \{\emptyset, X\}$.

9. Segue dalla definizione, dato che T preserva la misura.

□

Lemma 3.3.4. Se $\{a_n\} \subseteq \mathbb{R}$ è una successione che soddisfa $a_n \geq 0$ e $a_{n+m} \geq a_n + a_m$ allora esiste $\lim \frac{a_n}{n}$ e coincide con $\inf \frac{a_n}{n}$.

Dimostrazione. Fissato $m > 0$, per ogni $j > 0$ si consideri $j = km + n$ dove $0 \leq n < m$. Allora vale:

$$\frac{a_j}{j} = \frac{a_{n+km}}{n+km} \leq \frac{a_n}{km} + \frac{a_{km}}{km} \leq \frac{a_n}{km} + \frac{k \cdot a_m}{km} = \frac{a_n}{km} + \frac{a_m}{m}.$$

Ma siccome se $j \rightarrow \infty$ allora anche $k \rightarrow \infty$, si ottiene che

$$\limsup_j \frac{a_j}{j} \leq \frac{a_m}{m} \quad \text{e dunque} \quad \limsup_j \frac{a_j}{j} \leq \inf_m \frac{a_m}{m}.$$

Ma siccome $\inf a_m/m \leq \liminf a_j/j$ allora esiste

$$\lim_j \frac{a_j}{j} \quad \text{e coincide con} \quad \inf_j \frac{a_j}{j}.$$

□

Corollario 3.3.5. Se $\mathcal{A} \subseteq \mathcal{B}$ allora esiste sempre

$$\lim_{n \rightarrow \infty} \frac{1}{n} H \left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \right) = h(T, \mathcal{A}).$$

Dimostrazione. Basta prendere $a_n := H \left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \right)$ e mostrare che si può applicare il lemma precedente, ed infatti (usando le proprietà mostrate sopra):

$$\begin{aligned} a_{n+m} &= H \left(\bigvee_{i=0}^{n+m-1} T^{-i} \mathcal{A} \right) \leq H \left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \right) + H \left(\bigvee_{i=n}^{n+m-1} T^{-i} \mathcal{A} \right) = \\ &= a_n + H \left(\bigvee_{i=0}^{m-1} T^{-i} \mathcal{A} \right) = a_n + a_m. \end{aligned}$$

□

3.4 Alcune proprietà di $h(T, \mathcal{A})$ e di $h(T)$

Proposizione 3.4.1. *Se \mathcal{A}, \mathcal{C} sono sottoalgebre di \mathcal{B} e se T è una trasformazione che preserva la misura, allora valgono:*

1. $h(T, \mathcal{A}) \leq H(\mathcal{A})$.
2. $h(T, \mathcal{A} \vee \mathcal{C}) \leq h(T, \mathcal{A}) + h(T, \mathcal{C})$.
3. $\mathcal{A} \subseteq \mathcal{C} \implies h(T, \mathcal{A}) \leq h(T, \mathcal{C})$.
4. $h(T, \mathcal{A}) \leq h(T, \mathcal{C}) + H(\mathcal{A}|\mathcal{C})$.
5. Se T è invertibile e $m \geq 1$ allora $h(T, \mathcal{A}) = h\left(T, \bigvee_{i=0}^m T^i \mathcal{A}\right)$.

Dimostrazione.

1. Dalla (8) e (9) della proposizione 3.3.3 segue che:

$$\frac{1}{n} \cdot H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A}\right) \leq \frac{1}{n} \sum_{i=0}^{n-1} H(T^{-i} \mathcal{A}) = \frac{1}{n} \sum_{i=0}^{n-1} H(\mathcal{A}) = H(\mathcal{A}),$$

quindi la tesi si ottiene passando al limite su n .

2. Per le proprietà di H si ha che

$$\begin{aligned} H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{A} \vee \mathcal{C})\right) &= H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \vee \bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}\right) \leq \\ &\leq H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A}\right) + H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}\right). \end{aligned}$$

3. Se $\mathcal{A} \subseteq \mathcal{C}$ allora per ogni $n \geq 1$ si ha $\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \subseteq \bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}$. Quindi per ottenere la tesi basta usare la (4) della proposizione 3.3.3.
4. Usando la (4) e la (2) della proposizione 3.3.3 si ottiene

$$\begin{aligned} H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A}\right) &\leq H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \vee \bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}\right) = \\ &= H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}\right) + H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \middle| \bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}\right), \end{aligned}$$

ma per il secondo addendo, usando la (5) e la (9) della proposizione 3.3.3, si ottiene la disuguaglianza

$$\begin{aligned} H\left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \middle| \bigvee_{i=0}^{n-1} T^{-i} \mathcal{C}\right) &\leq \sum_{i=0}^{n-1} H\left(T^{-i} \mathcal{A} \middle| \left(\bigvee_{j=0}^{n-1} T^{-j} \mathcal{C}\right)\right) \leq \\ &\leq \sum_{i=0}^{n-1} H\left(T^{-i} \mathcal{A} \middle| T^{-i} \mathcal{C}\right) = n \cdot H(\mathcal{A}|\mathcal{C}). \end{aligned}$$

Dunque si ottiene

$$H\left(\bigvee T^{-i}\mathcal{A}\right) \leq H\left(\bigvee T^{-i}\mathcal{C}\right) + n \cdot H(\mathcal{A}|\mathcal{C})$$

e la tesi segue dalla definizione.

5. Usando la (3) e la (9) della proposizione 3.3.3 si ottiene

$$\begin{aligned} h\left(T, \bigvee_{i=-m}^m T^i\mathcal{A}\right) &= \lim_{k \rightarrow \infty} \frac{1}{k} \cdot H\left(\bigvee_{j=0}^{k-1} T^{-j}\left(\bigvee_{i=-m}^m T^i\mathcal{A}\right)\right) = \\ &= \lim_k \frac{1}{k} \cdot H\left(\bigvee_{i=-m}^{m+k-1} T^{-i}\mathcal{A}\right) = \lim_k \frac{1}{k} \cdot H\left(\bigvee_{i=-m}^m T^{-i}\mathcal{A} \vee \bigvee_{i=m}^{m+k-1} T^{-i}\mathcal{A}\right) = \\ &= \lim_k \frac{1}{k} \left[H\left(\bigvee_{i=-m}^{m+k-1} T^{-i}\mathcal{A}\right) + H\left(\bigvee_{-m}^m T^{-i}\mathcal{A} \middle| \bigvee_m^{m+k-1} T^{-i}\mathcal{A}\right) \right] = \\ &= \lim_k \frac{1}{k} \left[H\left(\bigvee_{i=0}^{k-1} T^{-i}\mathcal{A}\right) + H\left(\bigvee_{-m}^m T^{-i}\mathcal{A} \middle| \bigvee_m^{m+k-1} T^{-i}\mathcal{A}\right) \right]. \end{aligned}$$

Resta da mostrare che

$$\lim_{k \rightarrow \infty} \frac{1}{k} \cdot H\left(\bigvee_{-m}^m T^{-i}\mathcal{A} \middle| \bigvee_m^{m+k-1} T^{-i}\mathcal{A}\right) = 0,$$

ma questo, usando la (6) della proposizione 3.3.3, segue dal fatto che

$$H\left(\bigvee_{-m}^m T^{-i}\mathcal{A} \middle| \bigvee_m^{m+k-1} T^{-i}\mathcal{A}\right) \leq \frac{1}{k} \cdot H\left(\bigvee_{-m}^m T^{-i}\mathcal{A}\right).$$

□

Proposizione 3.4.2. *Se \mathcal{A} è una sottoalgebra finita di \mathcal{B} e se T è una trasformazione che preserva la misura, allora:*

$$h(T, \mathcal{A}) = \lim_{n \rightarrow \infty} H\left(\mathcal{A} \middle| \bigvee_{i=1}^n T^{-i}\mathcal{A}\right).$$

Dimostrazione. Innanzitutto, il limite esiste perché la successione considerata è debolmente decrescente (dalla (5) della proposizione 3.3.3). Per terminare la dimostrazione, si dimostra per induzione su $n \geq 2$ che (usando anche la (2) e la (9) della proposizione 3.3.3):

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{A}\right) = H(\mathcal{A}) + \sum_{j=1}^{n-1} H\left(\mathcal{A} \middle| \bigvee_{i=1}^j T^{-i}\mathcal{A}\right).$$

Infine, dividendo per n e passando al limite si ottiene la tesi perché il limite secondo Cesaro di una successione convergente coincide col limite ordinario. □

Proposizione 3.4.3. *Valgono le proprietà seguenti per $h(T)$:*

1. Per ogni $m \in \mathbb{N}^+$, $h(T^m) = m \cdot h(T)$.
2. Se T è invertibile allora, per ogni $m \in \mathbb{Z}$, $h(T^m) = |m| \cdot h(T)$.

Dimostrazione.

1. Innanzitutto si mostra che

$$h\left(T^m, \bigvee_{i=0}^{m-1} T^{-i} \mathcal{A}\right) = m \cdot h(T, \mathcal{A}).$$

Questo segue dal fatto che

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \cdot H\left(\bigvee_{j=0}^{k-1} T^{-mj} \left(\bigvee_{i=0}^{m-1} T^{-i} \mathcal{A}\right)\right) &= \\ &= \lim_{k \rightarrow \infty} \frac{m}{km} \cdot H\left(\bigvee_{i=0}^{km-1} T^{-i} \mathcal{A}\right) = m \cdot h(T, \mathcal{A}). \end{aligned}$$

Inoltre, indicando con \mathcal{A} una generica sotto- σ -algebra finita

$$\begin{aligned} m \cdot h(T) &= m \cdot \sup_{\mathcal{A}} h(T, \mathcal{A}) = \sup_{\mathcal{A}} h\left(T^m, \bigvee_{i=0}^{m-1} T^{-i} \mathcal{A}\right) \leq \\ &\leq \sup_{\mathcal{C}} h(T^m, \mathcal{C}) = h(T^m). \end{aligned}$$

Per ottenere la tesi resta da mostrare la disuguaglianza opposta, ed infatti

$$h(T^m, \mathcal{A}) \leq h\left(T^m, \bigvee_{i=0}^{m-1} T^{-i} \mathcal{A}\right) = m \cdot h(T, \mathcal{A}).$$

2. Per ottenere la tesi, basta dimostrare che $h(T^{-1}) = h(T)$; quindi, in definitiva, basta dimostrare che per ogni \mathcal{A} finita vale $h(T^{-1}, \mathcal{A}) = h(T, \mathcal{A})$. Ed infatti (per la (9) della proposizione 3.3.3)

$$H\left(\bigvee_{i=0}^{n-1} T^i \mathcal{A}\right) = H\left(T^{-(n-1)} \bigvee_{i=0}^{n-1} T^i \mathcal{A}\right) = H\left(\bigvee_{j=0}^{n-1} T^{-j} \mathcal{A}\right).$$

□

Lemma 3.4.4. *Sia (X, \mathcal{B}, m) uno spazio di probabilità e sia \mathcal{B}_0 un'algebra tale che la σ -algebra generata da \mathcal{B}_0 soddisfi $\sigma(\mathcal{B}_0) \doteq \mathcal{B}$. Se \mathcal{C} è una sottoalgebra finita di \mathcal{B} , allora per ogni $\varepsilon > 0$ esiste un'algebra finita $\mathcal{D} \subseteq \mathcal{B}_0$ tale che*

$$H(\mathcal{D}|\mathcal{C}) + H(\mathcal{C}|\mathcal{D}) < \varepsilon.$$

Dimostrazione. Si veda [1].

□

Corollario 3.4.5. *Se $\{\mathcal{A}_n\}_n$ è una successione crescente di algebre finite e se $\mathcal{C} \subseteq (\bigvee_n \mathcal{A}_n)$, allora*

$$\lim_n H(\mathcal{C}|\mathcal{A}_n) = 0.$$

Dimostrazione. Posto $\mathcal{B}_0 = \bigcup_{m \geq 1} \mathcal{A}_m$, allora questa è un'algebra e verifica che $\mathcal{C} \dot{\subset} \sigma(\mathcal{B}_0)$. Ma allora dal lemma precedente segue che per ogni $\varepsilon > 0$ esiste un'algebra finita $\mathcal{D}_\varepsilon \subseteq \mathcal{B}_0$ tale che $H(\mathcal{C}|\mathcal{D}_\varepsilon) < \varepsilon$. Ma siccome \mathcal{D}_ε è finita, allora per un opportuno $m_0 \in \mathbb{N}$ deve valere che $\mathcal{D}_\varepsilon \subseteq \mathcal{A}_{m_0}$, e quindi

$$\forall m \geq m_0 \quad H(\mathcal{C}|\mathcal{A}_m) \leq H(\mathcal{C}|\mathcal{A}_{m_0}) \leq H(\mathcal{C}|\mathcal{D}_\varepsilon) < \varepsilon ,$$

da cui segue la tesi. \square

Il seguente teorema risulta molto utile per il calcolo dell'entropia.

Teorema 3.4.6 (Kolmogorov-Sinai).

Sia $T : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ una trasformazione invertibile che preserva la misura e sia \mathcal{A} una sottoalgebra finita di \mathcal{B} tale che $(\bigvee_{n \in \mathbb{Z}} T^n \mathcal{A}) \doteq \mathcal{B}$. Allora

$$h(T) = h(T, \mathcal{A}) .$$

Dimostrazione. Siccome per definizione $h(T) = \sup h(T, \mathcal{D})$, dove l'estremo superiore è fatto al variare delle sottoalgebre finite $\mathcal{D} \subseteq \mathcal{B}$, allora per ottenere la tesi basta dimostrare che data una qualsiasi sottoalgebra finita \mathcal{C} di \mathcal{B} si ha che

$$h(T, \mathcal{C}) \leq h(T, \mathcal{A}) .$$

Ma infatti per le proprietà di $h(T, \cdot)$ si ha che

$$\begin{aligned} h(T, \mathcal{C}) &\leq h\left(T, \bigvee_{i=-m}^m T^i \mathcal{A}\right) + H\left(\mathcal{C} \left| \bigvee_{i=-m}^m T^i \mathcal{A}\right.\right) = \\ &= h(T, \mathcal{A}) + H\left(\mathcal{C} \left| \bigvee_{i=-m}^m T^i \mathcal{A}\right.\right) \end{aligned}$$

A questo punto la disuguaglianza voluta si ottiene ponendo $\mathcal{A}_m = \bigvee_{i=-m}^m T^i \mathcal{A}$ e applicando il corollario precedente, da cui si ottiene che

$$H(\mathcal{C}|\mathcal{A}_m) \xrightarrow{m \rightarrow +\infty} 0 .$$

\square

Proposizione 3.4.7. Se $T : X \rightarrow X$ è una trasformazione che preserva la misura (non necessariamente invertibile) e se \mathcal{A} è un'algebra finita contenuta in \mathcal{B} tale che $(\bigvee_{i \geq 0} T^{-i} \mathcal{A}) \doteq \mathcal{B}$, allora $h(T) = h(T, \mathcal{A})$.

Dimostrazione. La dimostrazione è analoga a quella del teorema precedente; basta sostituire $\bigvee_{i=0}^{m-1} T^{-i} \mathcal{A}$ al posto di $\bigvee_{i=-m}^m T^{-i} \mathcal{A}$ e usare la formula

$$h(T, \mathcal{A}) = h\left(T, \bigvee_{i=0}^{m-1} T^{-i} \mathcal{A}\right)$$

che si dimostra in modo analogo alla (5) della proposizione 3.4.1. \square

Alcuni risultati che risultano molto utili per il calcolo dell'entropia sono i seguenti.

Proposizione 3.4.8.

1. Se T è invertibile e $(\bigvee_{i \geq 0} T^{-i} \mathcal{A}) \doteq \mathcal{B}$ per qualche \mathcal{A} finita, allora $h(T) = 0$.
2. Se \mathcal{B}_0 è un'algebra tale che $\sigma(\mathcal{B}_0) \doteq \mathcal{B}$ allora indicando con \mathcal{A} una generica sotto- σ -algebra finita di \mathcal{B}_0

$$h(T) = \sup_{\mathcal{A}} h(T, \mathcal{A}) .$$

3. Sia $\{\mathcal{A}_n\}_n$ una successione crescente di sottoalgebra finite di \mathcal{B} tali che $(\bigvee_n \mathcal{A}_n) \doteq \mathcal{B}$. Allora: $h(T) = \lim_n h(T, \mathcal{A}_n)$.
4. Se $T_i : (X_i, \mathcal{B}_i, m_i) \rightarrow (X_i, \mathcal{B}_i, m_i)$ per $i = 1, 2$, allora vale la relazione: $h(T_1 \times T_2) = h(T_1) + h(T_2)$.

Dimostrazione.

1. Per la proposizione precedente e per la proposizione 3.4.2 si ha

$$h(T) = h(T, \mathcal{A}) = \lim_n H \left(\mathcal{A} \left| \bigvee_{i=1}^n T^{-i} \mathcal{A} \right. \right) .$$

Ma siccome $(\bigvee_{i \geq 1} T^{-i} \mathcal{A}) \doteq T^{-1} \mathcal{B} = \mathcal{B}$, allora posto $\mathcal{A}_n := (\bigvee_{i=1}^n T^{-i} \mathcal{A})$ si può applicare a tale successione il corollario 3.4.5 e si ottiene che $H(\mathcal{A} | \mathcal{A}_n) \rightarrow 0$ e quindi $h(T) = 0$.

2. Fissato $\varepsilon > 0$ e data $\mathcal{C} \subseteq \mathcal{B}$, per il lemma 3.4.4 esiste un'algebra finita $\mathcal{D}_\varepsilon \subset \mathcal{B}_0$ tale che $H(\mathcal{C} | \mathcal{D}_\varepsilon) < \varepsilon$. Inoltre, per la (4) della proposizione 3.4.1 si ha che

$$h(T, \mathcal{C}) \leq h(T, \mathcal{D}_\varepsilon) + H(\mathcal{C} | \mathcal{D}_\varepsilon) \leq h(T, \mathcal{D}_\varepsilon) + \varepsilon .$$

Quindi

$$h(T, \mathcal{C}) \leq \sup_{\substack{\mathcal{D} \subseteq \mathcal{B}_0 \\ \mathcal{D} \text{ finita}}} h(T, \mathcal{D}) + \varepsilon \quad \text{e infine:} \quad h(T) \leq \sup_{\substack{\mathcal{D} \subseteq \mathcal{B}_0 \\ \mathcal{D} \text{ finita}}} h(T, \mathcal{D}) .$$

La disuguaglianza opposta è ovvia per la definizione di $h(T)$.

3. Innanzitutto, il limite esiste perché $h(T, \mathcal{A}_n)$ è una successione crescente; inoltre $\mathcal{B}_0 := \bigcup_{n \geq 1} \mathcal{A}_n$ è un'algebra e vale che $\sigma(\mathcal{B}_0) = \mathcal{B}$. Per quanto dimostrato al punto precedente, si ha che

$$h(T) = \sup_{\substack{\mathcal{C} \subseteq \mathcal{B}_0 \\ \mathcal{C} \text{ finita}}} h(T, \mathcal{C}) .$$

A questo punto, se $\mathcal{C} \subseteq \mathcal{B}_0$ è finita, allora esiste un opportuno n_0 per cui $\mathcal{C} \subseteq \mathcal{A}_{n_0}$ e quindi $h(T, \mathcal{C}) \leq h(T, \mathcal{A}_{n_0})$, da cui segue

$$h(T) \leq \lim_{n \rightarrow \infty} h(T, \mathcal{A}_n) \quad , \text{ e la tesi:} \quad h(T) = \lim_{n \rightarrow \infty} h(T, \mathcal{A}_n) .$$

4. Se $\mathcal{A}_i \subseteq \mathcal{B}_i$ (con $i = 1, 2$) sono finite, allora anche $\mathcal{A}_1 \times \mathcal{A}_2$ è un'algebra finita e la partizione associata è

$$\xi(\mathcal{A}_1 \times \mathcal{A}_2) = \{A_1 \times A_2 \mid A_i \in \xi(\mathcal{A}_i)\} .$$

Ora, si consideri l'algebra \mathcal{T} formata dalle unioni finite di rettangoli misurabili; per definizione di σ -algebra prodotto si ha $\sigma(\mathcal{T}) = \mathcal{B}_1 \times \mathcal{B}_2$. Inoltre, dal punto (2) precedente segue che

$$h(T_1 \times T_2) = \sup_{\mathcal{C} \subset \mathcal{T} \text{ finita}} h(T_1 \times T_2, \mathcal{C}) .$$

Se $\mathcal{C} \subset \mathcal{T}$ è finita, allora esistono delle opportune algebre finite $\mathcal{A}_1 \subset \mathcal{B}_1$ e $\mathcal{A}_2 \subset \mathcal{B}_2$ tali che $\mathcal{C} \subset \mathcal{A}_1 \times \mathcal{A}_2$, quindi:

$$h(T_1 \times T_2) = \sup_{\substack{\mathcal{A}_1 \subset \mathcal{B}_1, \mathcal{A}_2 \subset \mathcal{B}_2 \\ \mathcal{A}_1, \mathcal{A}_2 \text{ finite}}} h(T_1 \times T_2, \mathcal{A}_1 \times \mathcal{A}_2) .$$

A questo punto, un rapido conto mostra che

$$\begin{aligned} H \left(\bigvee_{i=0}^{n-1} (T_1 \times T_2)^{-i}(\mathcal{A}_1 \times \mathcal{A}_2) \right) &= H \left(\left(\bigvee_{i=0}^{n-1} T_1^{-i} \mathcal{A}_1 \right) \times \left(\bigvee_{i=0}^{n-1} T_2^{-i} \mathcal{A}_2 \right) \right) = \\ &= - \sum \left[(m_1 \times m_2)(C_k \times D_j) \cdot \log(m_1 \times m_2)(C_k \times D_j) \right] = \\ &= H \left(\bigvee_{i=0}^{n-1} T_1^{-i} \mathcal{A}_1 \right) + H \left(\bigvee_{i=0}^{n-1} T_2^{-i} \mathcal{A}_2 \right) , \end{aligned}$$

dove i C_k, D_j sono rispettivamente elementi di $\xi \left(\bigvee_{i=0}^{n-1} T_h^{-i} \mathcal{A}_h \right)$ per $h = 1, 2$. Dunque $h(T_1 \times T_2, \mathcal{A}_1 \times \mathcal{A}_2) = h(T_1, \mathcal{A}_1) + h(T_2, \mathcal{A}_2)$ da cui segue la tesi: $h(T_1 \times T_2) = h(T_1) + h(T_2)$. □

3.5 Calcolo dell'entropia

Con i risultati ottenuti nella sezione precedente, possiamo procedere al calcolo effettivo dell'entropia in vari casi significativi.

Esempio 3.5.1. L'entropia dell'identità $id : (X, \mathcal{B}, m) \rightarrow (X, \mathcal{B}, m)$ è sempre nulla, perché $h(id, \mathcal{A}) = \lim_n \frac{1}{n} H(\mathcal{A}) = 0$.

Inoltre, se per qualche $p \in \mathbb{N}^+$ si ha $T^p = id$ allora $0 = h(T^p) = p \cdot h(T)$, da cui $h(T) = 0$.

Esempio 3.5.2. Se $T : S^1 \rightarrow S^1$ è la rotazione $T(z) = az$, con $a \in \mathbb{C}$, allora T ha entropia nulla. (In effetti, si può dimostrare che tale risultato vale in generale per ogni rotazione di un gruppo topologico compatto, abeliano e metrico). Ci sono due casi:

- Se a è una radice dell'unità, allora per un opportuno $p \in \mathbb{N}^+$ si ha $a^p = 1$ e quindi $T^p(z) = a^p z = z$, cioè $T = id$. Quindi per l'esempio precedente $h(T) = 0$.

- Se a non è una radice dell'unità, allora $\{a^n | n \in \mathbb{Z}\}$ è denso in S^1 e anche $\{a^n | n < 0\}$ lo è. A questo punto, si prendano i due semicerchi:

$$A_1 = S^1 \cap \{z \in \mathbb{C} | \Im(z) > 0\} \cup \{1\}, \quad A_2 = S^1 \cap \{z \in \mathbb{C} | \Im(z) < 0\} \cup \{-1\}.$$

e si definisca la partizione $\xi = \{A_1, A_2\}$. Allora per $n > 0$ si ha che $T^{-n}\xi = \{T^{-n}A_1, T^{-n}A_2\}$ consiste dei semicerchi che “iniziano” in a^{-n} oppure in $-a^{-n}$. Ma siccome $\{a^{-n} | n > 0\}$ è denso, allora ogni semicerchio deve appartenere a $\bigvee_{n \geq 0} T^{-n}\mathcal{A}(\xi)$. Dunque, ogni arco di S^1 appartiene a $\bigvee_{n \geq 0} T^{-n}\mathcal{A}(\xi)$, cioè quest'ultima coincide con la σ -algebra \mathcal{B} dei boreliani di S^1 . Ma allora per la (1) della proposizione 3.4.8 si ha $h(T) = 0$.

Esempio 3.5.3. Il (p_1, \dots, p_k) -shift T ha entropia $h(T) = -\sum_{i=1}^k p_i \cdot \log p_i$. In particolare, gli shift con vettori dei “pesi” dati da $(\frac{1}{2}, \frac{1}{2})$ e $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ hanno entropia rispettivamente data da $\log 2$ e da $\log 3$ e non possono essere trasformazioni coniugate.

Dimostrazione. (Nel caso di shift unilatero la dimostrazione è analoga). Sia $X = \prod_{-\infty}^{+\infty} \{1, \dots, k\}$ e per $i = 1, \dots, k$ si definisca $A_i = \{\{x_n\} \in X | x_0 = i\}$. Infine, si ponga $\xi = \{A_1, \dots, A_k\}$ e $\mathcal{A} = \mathcal{A}(\xi)$. A questo punto, per definizione della σ -algebra \mathcal{B} su X si ha

$$\bigvee_{i \in \mathbb{Z}} T^i \mathcal{A} = \mathcal{B}.$$

Per il teorema di Kolmogorov-Sinai 3.4.6 si ha

$$h(T) = h(T, \mathcal{A}) = \lim_n \frac{1}{n} \cdot H \left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \right).$$

Si osservi, ora, che un elemento generico della partizione $\xi \left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \right)$ è della forma

$$\begin{aligned} A_{i_0} \cap T^{-1} A_{i_1} \cap \dots \cap T^{-(n-1)} A_{i_{n-1}} &= \\ &= \left\{ \{x_n\} \in X \mid x_0 = i_0, x_1 = i_1, \dots, x_{n-1} = i_{n-1} \right\} \end{aligned}$$

ed ha misura $p_{i_0} \cdot p_{i_1} \cdot \dots \cdot p_{i_{n-1}}$. Dunque

$$\begin{aligned} H \left(\bigvee_{i=0}^{n-1} T^{-i} \mathcal{A} \right) &= - \sum (p_{i_0} \cdot \dots \cdot p_{i_{n-1}}) \cdot \log (p_{i_0} \cdot \dots \cdot p_{i_{n-1}}) = \\ &= - \sum_{i_0=\dots=i_{n-1}=1}^k (p_{i_0} \cdot \dots \cdot p_{i_{n-1}}) \cdot [\log p_{i_0} + \dots + \log p_{i_{n-1}}] = \\ &= -n \cdot \sum_{i=1}^k (p_i \cdot \log p_i), \end{aligned}$$

quindi si ottiene la tesi dalla definizione di $h(T, \mathcal{A})$. \square

Abbiamo osservato che nella definizione di entropia non è escluso il caso in cui essa assuma valore $+\infty$, a questo proposito si guardi l'esempio seguente.

Esempio 3.5.4. Sia $I = (0, 1]$ dotato della σ -algebra dei Boreliani e della misura di Lebesgue e sia $X = \prod_{-\infty}^{+\infty} I$ dotato della misura prodotto sulla σ -algebra prodotto. Se T è lo shift su X , allora $h(T) = +\infty$.

Per vedere ciò, si consideri

$$A_{n,i} = \left\{ \{x_j\} \in X \left| \frac{i-1}{n} < x_0 < \frac{i}{n} \right. \right\}, \text{ dove } n > 0, 1 \leq i \leq n$$

e si definisca una partizione di X come $\xi_n = \{A_{n,1}, \dots, A_{n,n}\}$. Allora si ha che $m(A_{n,i}) = 1/n$ e con un argomento analogo a quello usato nell'esempio precedente (e usando l'indipendenza di $\xi_n, T^{-1}\xi_n, \dots, T^{-k}\xi_n$) si ottiene che $h(T, \xi_n) = \log n$. Dunque, per ogni n si ha $h(T) \geq \log n$, cioè $h(T) = +\infty$.

3.6 β -espansioni

Introduzione

Sia $\beta > 1$ un numero reale e si consideri la trasformazione $T_\beta : [0, 1) \rightarrow [0, 1)$ definita da

$$T_\beta(x) := \beta x \pmod{1}.$$

Dato $x \in [0, 1)$ si può considerare $d_1 = d_1(x) := \lfloor \beta x \rfloor$ e in generale

$$d_n = d_n(x) := d_1(T_\beta^{n-1}x) = \lfloor \beta \cdot T_\beta^{n-1}x \rfloor.$$

Dal fatto che

$$T_\beta x = \beta x - d_1, \quad T_\beta^2 x = \beta T_\beta x - d_2, \quad \dots$$

segue che

$$x = \frac{d_1}{\beta} + \frac{T_\beta x}{\beta} = \frac{d_1}{\beta} + \frac{d_2}{\beta} + \dots$$

e quindi si possono definire le **cifre della β -espansione** di x come sopra $d_n = \lfloor \beta T_\beta^{n-1}x \rfloor$ (con $n \geq 1$) e chiaramente $d_n \in \{0, 1, \dots, \lfloor \beta \rfloor\}$.

La trasformazione T_β , definita sopra, genera lo sviluppo in serie

$$x = \sum_{k=1}^{+\infty} \frac{d_k}{\beta^k}$$

dove $d_k = d_k(x) = \lfloor \beta \cdot T_\beta^{k-1}x \rfloor$, per $k \geq 1$. Tale β -espansione verrà indicata da $0.d_1d_2\dots$ per sottolineare l'analogia con l'espansione decimale.

Osservazione 3.6.1. Si osservi che nel caso in cui $\beta = 10$, allora lo sviluppo decimale usuale coincide con la 10-espansione. Infatti in tal caso $T_{10}x = 10x \pmod{1}$ e se $x \in [0, 1)$ ha uno sviluppo in cifre decimali dato da $x = 0.a_1a_2\dots$ allora si ha:

$$\begin{aligned} a_1 &= \lfloor 10 \cdot 0.a_1a_2\dots \rfloor = \lfloor 10x \rfloor = d_1(x) \\ a_2 &= \lfloor 10 \cdot 0.a_2\dots \rfloor = \lfloor 10 \cdot T_{10}x \rfloor = d_2(x) \\ \vdots & \quad \quad \quad \vdots \end{aligned}$$

Se si considera la misura di Lebesgue λ su $[0, 1)$ allora T_β non preserva tale misura se $\beta \notin \mathbb{N}$.

Infatti, presi $a, b \in [0, 1)$ tali che $\beta - \lfloor \beta \rfloor < a < b < 1$, allora

$$T_\beta^{-1}(a, b) = \bigcup_{j=0}^{\lfloor \beta \rfloor - 1} \left(\frac{a}{\beta} + \frac{j}{\beta}, \frac{b}{\beta} + \frac{j}{\beta} \right)$$

e dunque

$$\lambda(T_\beta^{-1}(a, b)) = \sum_{j=0}^{\lfloor \beta \rfloor - 1} \frac{b-a}{\beta} = \frac{\lfloor \beta \rfloor}{\beta} (b-a).$$

Ma a questo punto, siccome $\beta \notin \mathbb{N}$ allora $\lfloor \beta \rfloor / \beta < 1$ e quindi $\lambda(T_\beta^{-1}(a, b)) < \lambda(a, b)$.

Un problema naturale allora è il seguente: *esiste un'opportuna misura ν_β per cui T_β sia ν_β -invariante e della forma $\nu_\beta(A) = \int_A g_\beta(x) d\lambda(x)$, dove g_β è una funzione che soddisfa $0 < g_\beta < \infty$?* Si noti che se tale misura ν esiste, allora è una misura equivalente a quella di Lebesgue, ovvero ν e λ hanno gli stessi insiemi di misura nulla.

Una risposta a tale problema è stata data da A.O.Gelfond nel 1959 e precisamente

$$\nu_\beta(A) = \int_A g_\beta(x) d\lambda(x),$$

dove g_β è una funzione λ -misurabile che verifica

$$1 - \frac{1}{\beta} \leq g_\beta(x) \leq \frac{1}{1 - \frac{1}{\beta}}.$$

definita da

$$g_\beta(x) = \frac{1}{F(\beta)} \sum_{n \in \mathbb{N}: x < T_\beta^n(1)} \frac{1}{\beta^n}, \quad x \in [0, 1),$$

con costante di normalizzazione

$$F(\beta) = \int_0^1 \left(\sum_{n \in \mathbb{N}: x < T_\beta^n(1)} \frac{1}{\beta^n} \right) d\lambda(x).$$

Proprietà ergodiche ed entropia

Lemma 3.6.2 (Knopp).

Sia $B \subseteq [0, 1)$ un insieme Lebesgue-misurabile e sia \mathcal{C} una famiglia di sotto-intervalli di $[0, 1)$ tale che:

1. ogni sotto-intervallo aperto di $[0, 1)$ è unione al più numerabile di elementi disgiunti di \mathcal{C} ;
2. esiste $\gamma > 0$ tale che per ogni $A \in \mathcal{C}$ si ha che $\lambda(A \cap B) \geq \gamma \lambda(A)$.

Allora $\lambda(B) = 1$.

Dimostrazione. Si veda [3]. □

Definizione 3.6.3. *L'insieme definito da*

$$\Delta(i_1, \dots, i_n) = \{x \in [0, 1) \mid d_1(x) = i_1, \dots, d_n(x) = i_n\},$$

dove $x = \sum_{k=1}^{\infty} \frac{d_k(x)}{\beta^k}$ è la β -espansione di x , è detto **intervallo fondamentale (o cilindro)** di ordine n . Un tale $\Delta(i_1, \dots, i_n)$ è detto **intervallo pieno** di ordine n se

$$\lambda(T_\beta^n(\Delta(i_1, \dots, i_n))) = 1,$$

altrimenti è detto **non-pieno**.

Si può vedere che ogni intervallo pieno di ordine n ha misura $1/\beta^n$. Inoltre, si noti che gli intervalli non-pieni di ordine n sono esattamente quelli per cui l'ultima cifra i_n è uguale a 1 e hanno misura uguale a $1/\beta^{n+1}$.

Lemma 3.6.4.

1. Dato $n \geq 1$, si può ricoprire $[0, 1)$ con intervalli pieni disgiunti di ordine n o $n+1$.
2. (Corollario di 1) Ogni intervallo è unione al più numerabile di intervalli pieni.

Dimostrazione. Si veda [3]. □

Proposizione 3.6.5. *Se $\beta > 1$, allora la mappa $T_\beta(x) = \beta x \pmod{1}$ è ergodica.*

Dimostrazione. Sia $B \subset [0, 1)$ un insieme λ -misurabile e T_β -invariante con misura positiva, allora è sufficiente dimostrare che $\lambda(B) = 1$. Sia \mathcal{C} la famiglia di tutti gli intervalli pieni, allora essa verifica l'ipotesi (1) del lemma di Knopp 3.6.2. Ora, sia $E \in \mathcal{C}$ un intervallo pieno di ordine n , allora per ogni insieme Lebesgue-misurabile C si ha $\lambda(T^{-n}\beta C \cap E) = \beta^{-n}\lambda(C)$ e quindi

$$\frac{\lambda(B \cap E)}{\lambda(E)} = \frac{\lambda(T^{-n}\beta B \cap E)}{\lambda(E)} = \frac{\beta^{-n}\lambda(B)}{\beta^{-n}} = \lambda(B),$$

da cui segue che per ogni intervallo pieno E di ordine n vale $\lambda(B \cap E) = \lambda(B)\lambda(E)$. Infine, applicando il lemma di Knopp con $\gamma = \lambda(B)$ si ottiene che $\lambda(B) = 1$. □

Proposizione 3.6.6. *Per ogni $\beta > 1$, $h(T_\beta) = \log \beta$.*

Dimostrazione. Si veda [3]. □

Bibliografia

- [1] Peter Walters. *Ergodic Theory - Introductory Lecturers*. Springer-Verlag, 1943.
- [2] Paul R. Halmos. *Lecturers on Ergodic Theory*. The Mathematical Society of Japan, 1956.
- [3] Karma Dajani, Cor Kraaikamp. *Ergodic Theory of Numbers*. The Mathematical Association of America, 2002.
- [4] James R. Brown. *Ergodic Theory and Topological Dynamics*. Academic Press, 1976.