CAPITOLO 1

I concetti fondamentali

1. Introduzione

Non è possibile introdurre la Teoria Ergodica senza fare cenno alle sue origini fisiche.

La Teoria Ergodica rappresenta uno dei tentativi di spiegare le proprietà termodinamiche dei sistemi macroscopici, vale a dire dei sistemi composti da un numero di particelle almeno pari al numero di AVOGADRO, 6×10^{23}. La disciplina che vuole di fatto dimostrare i principi della termodinamica come teoremi a partire dalle equazioni della dinamica è la *Meccanica Statistica.* In particolare, la Meccanica Statistica vuole spiegare come il comportamento irreversibile dei sistemi termodinamici, si pensi al secondo principio della termodinamica, possa derivare dalle equazioni della dinamica che sono reversibili. La letteratura scientifica dedicata alla meccanica statistica è veramente sterminata. Qui citiamo solo, per la termodinamica due classici [**16, 55**] e [**74**], per la teoria cinetica ed il cosiddetto teorema H di BOLTZMANN [**12**], per la meccanica statistica in generale [**31, 62, 69**]. Piú specificamente per i fondamenti della meccanica statistica si può consultare il libro di GIBBS [**22**] come un modo di porre i fondamenti in maniera assai distante da quella proposta dalla teoria ergodica. Per l'approccio alla meccanica statistica attraverso la teoria ergodica si vedano [**37**] e [**15**]. Anche i fisici che preferiscono l'approccio alla meccanica statistica basato sulla teoria degli *ensembles* introdotto da GIBBS e EINSTEIN, si veda, per esempio [**23**], ritengono tuttavia che la teoria ergodica abbia fornito importanti contributi ai fondamenti della teoria fisica.

Ricordiamo che, nella meccanica classica, lo stato di un sistema con n gradi di libertà è descritto dalle n coordinate generalizzate $q = (q_1, \ldots, q_n)$ e dai loro momenti coniugati $p = (p_1, p_2, \ldots, p_n)$. Ogni stato di tale sistema è cosí rappresentato da un punto P di coordinate (q, p) in un sottoinsieme di $\mathbf{R}^{2n}$ detto *spazio delle fasi.* Il moto di tale punto rappresentativo, e, quindi, del sistema considerato, è retto dalle equazioni di HAMILTON

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}, \qquad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j} \qquad (j = 1, 2, \dots, n).$$

Qui H è la (funzione) Hamiltoniana del sistema, che è indipendente dal tempo t e, di solito, si esprime come la somma dell'energia cinetica $K(p)$ e dell'energia potenziale $V(q)$. Detto $P_0 = (q_0, p_0)$ lo stato iniziale del sistema, la soluzione delle equazioni di HAMILTON determina lo stato del sistema per istante $t > 0$

$$(q_t, p_t) = T_t\ (q_0, p_0)\,.$$

Si ottiene pertanto un semigruppo $\{T_t : t \geq 0\}$ di trasformazioni,

$$T_{t+t'} = T_t\, T_{t'},$$

sullo spazio delle fasi. Tale semigruppo di trasformazioni si dice *flusso hamiltoniano* e gode della proprietà fondamentale espressa dal seguente teorema.

TEOREMA 1.1 (di Liouville). *Il flusso hamiltoniano $\{T_t\}$ conserva la misura di* LEBESGUE *λ_{2n} nello spazio delle fasi: se A è un sottoinsieme misurabile dello spazio delle fasi si ha, per ogni $t \geq 0$,*

$$\lambda_{2n}(A) = \lambda_{2n}\,(T_t\, A)\,.$$

Nel seguito ci occuperemo esclusivamente di trasformazioni che conservano la misura, adottando però uno schema nei quali il tempo è discreto.

Se $T : \Omega \to \Omega$ è una trasformazione che conserva la misura, l'orbita $\{T^n\omega : n \in \mathbf{Z}\}$ del punto $\omega \in \Omega$ rappresenta la storia completa di tale punto. Se f è una grandezza fisica, passibile di misurazione, allora

$$f(\omega),\ \ f(T\,\omega), \dots, f\,(T^n\,\omega)\,, \dots$$

sono i valori di f in istanti successivi. Nella meccanica statisitca di un sistema composto da piú (molte) particelle nessuna misurazione è istantanea (si pensi solo alle collisioni molecolari) sicché ogni misurazione non riguarda i valori istantanei, bensí una media temporale

$$\frac{1}{n}\sum_{j=0}^{n-1} f\,(T^n\,\omega)\,,$$

che è solitamente molto difficile, se non impossibile, da calcolare a causa della mancanza della soluzione esplicita delle equazioni di HAMILTON. L'ipotesi ergodica consite nella possibilità di sostituire la media temporale

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1} f\,(T^n\,\omega)\,,$$

supposto che tale limite esista, con la media spaziale

$$\int_\Omega f \, d\mu,$$

per la quale non occorrono le soluzioni dele equazioni del moto, ma basta la conoscenza della “geometria” del sistema.

2. Trasformazioni che conservano la misura

DEFINIZIONE 2.1. Dati due spazî mensurali $(\Omega_1, \mathcal{F}_1, \mu_1)$ e $(\Omega_2, \mathcal{F}_2, \mu_2)$, si dice che una trasformazione $T : \Omega_1 \to \Omega_2$

(i) è *misurabile* se appartiene alla tribú $\mathcal{F}_1$ l'immagine inversa mediante T di ogni insieme di $\mathcal{F}_2$, vale a dire se è $T^{-1}\mathcal{F}_2 \subset \mathcal{F}_1$,oppure, equivalentemente, se $T^{-1}A \in \mathcal{F}_1$ per ogni insieme A di $\mathcal{F}_2$;
(ii) *conserva la misura* se T è misurabile e $\mu_1\left(T^{-1}A\right) = \mu_2(A)$ per ogni insieme $A \in \mathcal{F}_2$.

Spesso si ha la situazione $\Omega_1 = \Omega_2 = \Omega$, si considerano cioè trasformazioni di uno spazio Ω in sé; inoltre spesso tanto μ_1 quanto μ_2 sono misure di probabilità.

Per indicare una trasformazione che conserva la misura si potrebbe ricorrere alla notazione, assai pesante, $T : (\Omega_1, \mathcal{F}_1, \mu_1) \longrightarrow (\Omega_2, \mathcal{F}_2, \mu_2)$; tuttavia, ove non sia possibile confusione tra le tribú e le misure in gioco, si scriverà semplicemente $T : \Omega_1 \to \Omega_2$.

Se $T : \Omega_1 \to \Omega_2$ e $T' : \Omega_2 \to \Omega_3$ conservano la misura (tra i rispettivi spazî), altrettanto fa la trasformazione composta $S := T' \circ T : \Omega_1 \to \Omega_3$ tra gli spazî $(\Omega_1, \mathcal{F}_1, \mu_1)$ e $(\Omega_3, \mathcal{F}_3, \mu_3)$; infatti, se A è un insieme di $\mathcal{F}_3$ si ha

$$\begin{aligned}\mu_1\left(S^{-1}A\right) &= \mu_1\left[(T' \circ T)^{-1}\, A\right] = \mu_1\left[\left(T^{-1} \circ T'^{-1}\right)\, A\right] \\ &= \mu_1\left[T^{-1}\left(T'^{-1}\, A\right)\right] = \mu_2\left(T'^{-1}\, A\right) = \mu_3(A).\end{aligned}$$

La definizione 2.1 è in pratica di applicazione piuttosto difficile. Sono perciò utili i due teoremi che seguono e ai quali occorrerà premettere un richiamo. Ricordiamo che si chiama *semi–anello* una famiglia $\mathcal{S}$ di sottoinsiemi di un insieme non vuoto Ω tale che

(i) $\emptyset \in \mathcal{S}$;
(ii) $A, B \in \mathcal{S} \implies A \cap B \in \mathcal{S}$ ($\mathcal{S}$ è stabile rispetto all'intersezione);

(iii) per ogni coppia A e B di insiemi di $\mathcal{S}$ esistono insiemi disgiunti

$$E_1, E_2, \ldots, E_n$$

di $\mathcal{S}$ tali che

$$A \setminus B = \cup_{j=1}^n E_j.$$

Ricordiamo (si veda, per esempio, [**38**, Theorem 1.4], che l'anello $\mathbf{R}(\mathcal{S})$ generato dal semi–anello $\mathcal{S}$ è composto da tutti, e soli, gli insiemi E che possono essere espressi come l'unione

$$E = \cup_{j=1}^n A_j \tag{2.1}$$

di un numero finito di insiemi disgiunti di $\mathcal{S}$.

TEOREMA 2.1. *Siano* $(\Omega_1, \mathcal{F}_1, \mu_1)$ *e* $(\Omega_2, \mathcal{F}_2, \mu_2)$ *spazî mensurali, sia* $\mathcal{S}_2$ *un semi–anello contenuto in* $\mathcal{F}_2$ *e sia* $T : \Omega_1 \to \Omega_2$ *una trasformazione tale che, per ogni insieme appartenente al semi–anello* $\mathcal{S}_2$, $T^{-1} E$ *appartenga alla tribú* $\mathcal{F}_1$ *e tale che*

$$\mu_1\left(T^{-1} E\right) = \mu_2(E). \tag{2.2}$$

Allora la (2.2) *vale per tutti gli insiemi dell'anello* $\mathbf{R}_2 = \mathbf{R}_2(\mathcal{S}_2)$ *generato da* $\mathcal{S}_2$.

DIMOSTRAZIONE. In virtú della (2.1), ogni insieme A di $\mathbf{R}_2$ si può esprimere come unione disgiunta di un numero finito di insiemi di $\mathcal{S}$,

$$A = \cup_{j=1}^n E_j.$$

Pertanto

$$\begin{aligned}\mu_1\left(T^{-1} A\right) &= \mu_1\left[T^{-1}\left(\cup_{j=1}^n E_j\right)\right] = \mu_1\left(\cup_{j=1}^n T^{-1} E_j\right) = \sum_{j=1}^n \mu_1\left(T^{-1} E_j\right) \\ &= \sum_{j=1}^n \mu_2(E_j) = \mu_2\left(\cup_{j=1}^n E_j\right) = \mu_2(A),\end{aligned}$$

che stabilisce l'asserto. □

TEOREMA 2.2. *Siano* $(\Omega_1, \mathcal{F}_1, \mu_1)$ *e* $(\Omega_2, \mathcal{F}_2, \mu_2)$ *spazî mensurali, sia* $\mathcal{A}_2$ *un'algebra che genera la tribú* $\mathcal{F}_2$, $\mathcal{F}_2 = \mathcal{F}(\mathcal{A}_2)$ *e sia* $T : \Omega_1 \to \Omega_2$ *una trasformazione tale che, per ogni insieme* B *di* $\mathcal{A}_2$,

(i) $T^{-1} B \in \mathcal{F}_1$
(ii) $\mu_1\left(T^{-1} B\right) = \mu_2(B)$.

Allora T *conserva la misura.*

DIMOSTRAZIONE. Si ponga

$$\mathcal{C} := \left\{A \in \mathcal{F}_2 : T^{-1} A \in \mathcal{F}_1, \mu_1\left(T^{-1} A\right) = \mu_1(A)\right\}$$

Da questa posizione segue che valgono le inclusioni $\mathcal{A}_2 \subset \mathcal{C} \subset \mathcal{F}_2$. Per dimostrare che $\mathcal{C} = \mathcal{F}_2$, basterà quindi dimostrare che vale l'inclusione $\mathcal{F}_2 \subset \mathcal{C}$. La dimostrazione sarà conclusa quando si sarà dimostrato che $\mathcal{C}$ è una classe monotona; infatti, il teorema della classe monotona (si veda, per esempio, [**25**, Theorem 6-B]) assicura che sia $\mathcal{F}_2 \subset \mathcal{C}$. □

3. Esempî

ESEMPIO 3.1. La trasformazione identità $I : \Omega \longrightarrow \Omega$, $I\,\omega = \omega$, per ogni $\omega \in \Omega$, ovvimante conserva la misura sullo spazio $(\Omega, \mathcal{F}, \mu)$.

Prima di dare i prossimi esempî occorre richiamere le misura di HAAR. Per le misura di HAAR si veda [**48**].

Sia G un gruppo topologico compatto; esiste allora una misura finita μ definita sui boreliani $\mathcal{B}(G)$ di G tale che, per ogni $x \in G$ e per ogni $A \in \mathcal{B}(G)$, sia

$$\mu(x\,A) = \mu(A).$$

Se ν e μ sono due misure di HAAR definite su $\mathcal{B}(G)$, allora esiste una costante $c > 0$ tale che $\nu = c\,\mu$. Esiste, pertanto, un'unica misura di probabilità di HAAR definita sullo spazio misurabile $(\mathcal{G}, \mathcal{B}(G))$.

ESEMPIO 3.2. Sia $\mathcal{K}$ la circonferenza unitaria nel piano complesso, $\mathcal{K} := \{z \in \mathbf{C} : |z| = 1\}$, la si doti della tribú di BOREL $\mathcal{B}$ e sia μ la misura di HAAR (normalizzata) su $\mathcal{B}$. Fissato un elemento a di $\mathcal{K}$, si definisca $T_a : \mathcal{K} \longrightarrow \mathcal{K}$ mediante $T_a\,z := a\,z$. poiché μ è la misura di HAAR si ha

$$\mu\left(T_a^{-1} A\right) = \mu(A)$$

per ogni boreliano A.

ESEMPIO 3.3. (del quale l'esempio precedente è un caso particolare). Siano G un gruppo topologico compatto ed a un elemento di G. Allora $T_a : G \longrightarrow G$ definita da $T_a\,z := a\,z$ conserva la misura di HAAR (normalizzata).

ESEMPIO 3.4. (del quale l'esempio precedente è un caso particolare). Siano G un gruppo topologico compatto e $T : G \longrightarrow G$ un endomorfismo suriettivo. Allora T preserva la misura di HAAR normalizzata μ. Infatti, si definisca sui boreliani $\mathcal{B}(G)$ di G la misura di probabilità

$$\nu(A) := \mu\left(T^{-1} A\right) \qquad (A \in \mathcal{B}(G)).$$

Quale che sia $x \in G$ si ha, per $E \in \mathcal{B}(G)$,

$$\nu(T x \cdot E) = \mu\left[T^{-1}\left(T x \cdot E\right)\right] = \mu\left(x \cdot T^{-1} E\right) = \mu\left(T^{-1} E\right) = \nu(E).$$

Ma T è suriettiva e perciò l'unicità della misura di HAAR dà l'asserto.

ESEMPIO 3.5. Sia G un gruppo topologico compatto. Una trasformazione *affine* A su G è definita da $A(x) := a \cdot T(x)$ ove a è un elemento dato di G e $T : G \longrightarrow G$ è un endomorfismo suriettivo di G in sé. La trasformazione A, come composizione di due trasformazioni che conservano la misura di HAAR conserva la misura di HAAR.

ESEMPIO 3.6. (La traslazione, *shift*, bilatera). Si consideri l'insieme finito $X := \{0, 1, \dots, k-1\}$ e sia $p = (p_0, p_1, \dots, p_{k-1})$ una legge di probabilità. Naturalmente, qui e nel seguito supporremo, senza richiamarlo ogni volta, che sia $p_j > 0$ per ogni indice $j \in \{0, 1, \dots, k-1\}$. Sia

$$\Omega := \prod_{n \in \mathbf{Z}} X$$

l'insieme $X^{\mathbf{Z}}$ di tutte le successioni doppiamente infinite di elementi di X,

$$\Omega = \{(\dots, x_{-n}, \dots, x_{-1}, x_0, x_1, \dots, x_n, \dots) : x_n \in X \ (n \in \mathbf{Z})\},$$

munito della tribú prodotto

$$\mathcal{B} := \dots \otimes \mathcal{P}(X) \otimes \mathcal{P}(X) \otimes \dots \otimes \mathcal{P}(X) \otimes \dots .$$

Si consideri ora la misura prodotto μ definita sullo spazio misurabile $(\Omega, \mathcal{B})$. Si definisca la trasformazione $T : \Omega \longrightarrow \Omega$ mediante

$$T(\{x_n\}) = \{y_n : n \in \mathbf{Z}\},$$

ove, per ogni $n \in \mathbf{Z}$, $y_n := x_{n+1}$. La trasformazione T cosí definita conserva la misura di tutti i cilindri misurabili e, quindi, di tutti le unioni finite e disgiunte di cilindri misurabili; l'insieme di tali unioni costituisce un'algebra sicché il Teorema 2.2 assicura che T conservi la misura.

ESEMPIO 3.7. Siano X e p come nell'esempio precedente e sia

$$\Omega := \prod_{n \in \mathbf{Z}_+} X$$

l'insieme X^{Z_+} di tutte le successioni di elementi di X

$$\Omega = \left\{(x_0, x_1, \dots, x_n, \dots) : x_n \in X \ (n \in \mathbf{Z})\right\},$$

munito della tribú prodotto

$$\mathcal{B} := \mathcal{P}(X) \otimes \mathcal{P}(X) \otimes \dots \otimes \mathcal{P}(X) \otimes \dots .$$

Si consideri ora la misura prodotto μ definita sullo spazio misurabile $(\Omega, \mathcal{B})$. Di nuovo, si definisca la trasformazione $T : \Omega \longrightarrow \Omega$ mediante

$$T(\{x_n\}) = \{y_n : n \in \mathbf{Z}\},$$

ove, per ogni $n \in \mathbf{Z}_+$, $y_n := x_{n+1}$. Lo stesso ragionamento dell'esempio precedente mostra che T conserva la misura; si osservi però, che mentre la traslazione bilatera è invertibile, tale non è la traslazione unilatera. Si osservi, inoltre, che se l'insieme misurabile A è invariante, $A = T^{-1} A$, rispetto alla traslazione unilatera T, allora si ha $A = T A$.

Infatti, sia $\omega \in T A$; allora, $T^{-1} \omega \in A = T^{-1} A$. Per definizione di T^{-1} si ha, quindi, $T T^{-1} \omega \in A$. Ma $T T^{-1} = I$ (l'operatore identità) sicché ω appartiene ad A. Dunque, $T A \subset A$. Ora sia $\omega \in A$, ovvero, $\omega = T T^{-1} \omega \in A$, sicché $T^{-1} \omega \in T^{-1} A = A$, cioè $T^{-1} \omega$ appartiene ad $T^{-1} A = A$, o, ancora, $\omega \in TA$ che significa $A \subset T A$. Le due inclusioni danno l'asserto.

Il viceversa non è vero, nel senso che esistono insiemi A per i quali $A = T A$ senza che essi siano invarianti. Un tale insieme è dato da

$$A := \{(x, x, \dots, x, \dots) : x \in X\} .$$

In questo caso si ha, per ogi $x \in X$, $T(x, x, \dots, x, \dots) = (x, x, \dots, x, \dots)$, vale a dire $T A = A$. D'altro canto si ha anche

$$T^{-1}(x, x, \dots, x, \dots) = \cup_{y \in X} \{(y, x, x, \dots, x, \dots)\}.$$

Perciò, $T^{-1} A \supset A$ ma $A \neq T^{-1} A$.

Esempio 3.8. (La traslazione, *shift*, markoviana bilatera). Si ritorni all'esempio 3.6 e si consideri come lí l'insieme $X := \{0, 1, \dots, k-1\}$, il prodotto

$$\Omega := \prod_{n \in \mathbf{Z}} X,$$

munito della tribú prodotto $\mathcal{B} := \dots \otimes \mathcal{P}(X) \otimes \mathcal{P}(X) \otimes \dots \otimes \mathcal{P}(X) \otimes \dots$ e la traslazione bilatera

$$T\{x_n\} := \{y_n\},$$

ove, per ogni $n \in \mathbf{Z}$, $y_n := x_{n+1}$.

Si supponga ora che, per ogni $n \in \mathbf{Z}$, sia assegnata una funzione $p_n : X^{n+1} \longrightarrow [0, 1]$ tale che

$$\sum_{i=0}^{k-1} p_0(i) = 1 \tag{a}$$

$$p_n\big(i(0), i(1), \dots, i(n)\big) = \sum_{i=0}^{k-1} p_{n+1}\big(i(0), i(1), \dots, i(n), i\big). \tag{b}$$

Sul semianello $\mathcal{S}_e$ dei cilindri elementari si definisca una funzione d'insieme $\mu : \mathcal{S}_e \longrightarrow \mathbf{R}_+$ mediante

$$\mu\left(\big\{\{x_n\} : x_s = i(0), x_{s+1} = i(1), \dots, x_{s+n} = i(n)\big\}\right) := p_n\big(i(0), i(1), \dots, i(n)\big).$$

Allora il teorema di coerenza di KOLMOGOROV assicura che μ possa essere estesa ad una misura di probabilità su $(\Omega, \mathcal{B})$, che indicheremo ancora con μ, e che, per i Teoremi 2.1 e 2.2, è conservata dalla trasformazione T.

Si riottiene l'Esempio 3.6 prendendo

$$p_n\,(a_0, a_1, \dots, a_n) = p_{a_0}\, p_{a_1}\, \dots\, p_{a_n}.$$

Siano ora Π la matrice di transizione $k \times k$ di una catena di MARKOV che ha X come insieme degli stati e $p = (p_0, p_1, \dots, p_{k-1})$ un vettore (riga) di probabilità e si ponga

$$p_n\big(i(0), i(1), \dots, i(n)\big) := p_{i(0)}\, p_{i(0)\,i(1)}\, p_{i(1)\,i(2)}\, \cdots\, p_{i(n-1)\,i(n)}.$$

Con questa scelta della funzione p_n la trasformazione T, che, come si è visto sopra, conserva la misura, si dice *traslazione, shift, bilatera di* MARKOV–(p, Π). La traslazione bilatera dell'Esempio 3.6 è una traslazione di MARKOV se si prende $p = (p_0, p_1, \dots, p_{k-1})$ e $\Pi = (p_{ij})$ con $p_{ij} = p_j$ quale che sia $i \in \{0, 1, \dots, k-1\}$.

La trasformazione di MARKOV unilatera si costruisce come sopra effettuando gli ovvî cambiamenti.

ESEMPIO 3.9. Sia dato il quadrato unitario $[0,1]^2$, dotato della tribú di BOREL $\mathcal{B}([0,1]^2)$ e della restrizione λ_2 della misura di LEBESGUE a questa; allora, la proiezione canonica $\pi(x, y) := x$ conserva la misura tra gli spazî $\left([0,1]^2, \mathcal{B}([0,1]^2), \lambda_2\right)$ e $([0,1], \mathcal{B}([0,1]), \lambda)$.

OSSERVAZIONE 3.1. Siano Ω_1 un arbitrario insieme non vuoto, $(\Omega_2, \mathcal{F}_2, \mu_2)$ un qualsiasi spazio di probabilità e $T : \Omega_1 \longrightarrow \Omega_2$ un'arbitraria trasfomazione suriettiva; si possono, allora, scegliere una tribú $\mathcal{F}_1$ di sottoinsiemi di Ω_1 ed una misura di probabilità μ_1 su $\mathcal{F}_1$ in modo che T conservi la misura tra $(\Omega_1, \mathcal{F}_1, \mu_1)$ e $(\Omega_2, \mathcal{F}_2, \mu_2)$. Basta porre

$$\mathcal{F}_1 := \left\{T^{-1}\, B : B \in \mathcal{F}_2\right\}$$

e definire μ_1 mediante

$$\mu_1\left(T^{-1}\, B\right) := \mu_2(B).$$

Viceversa, se $(\Omega_1, \mathcal{F}_1, \mu_1)$ è uno spazio di probabilità qualsiasi, se Ω_2 è un arbitrario insieme non vuoto e $T : \Omega_1 \longrightarrow \Omega_2$ un'arbitraria trasformazione suriettiva, si possono scegliere una tribú $\mathcal{F}_2$ di sottoinsiemi di Ω_2 ed una misura di probabilità μ_2 su $\mathcal{F}_2$ in modo che T conservi la misura. Basta prendere

$$\mathcal{F}_2 := \left\{B \subset \Omega_2 : T^{-1}\, B \in \mathcal{F}_1\right\}$$

e definire μ_2 mediante

$$\mu_2(B) := \mu_1\left(T^{-1}\, B\right).$$

ESEMPIO 3.10. Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio di probabilità e su questo si consideri il processo stocastico a tempo discreto $\{X_n : n \in \mathbf{Z}\}$. Si supponga che tale processo sia stazionario, nel senso che, per ogni scelta di un numero naturale $k \in \mathbf{N}$, di k numeri interi $n_1, n_2, \dots, n_k$, di k boreliani $B_1, B_2, \dots, B_k$ e di un numero intero $n \in \mathbf{Z}$ si abbia

$$\mu\left(\cap_{j=1}^{k} X_{n_j}^{-1}(B_j)\right) = \mu\left(\cap_{j=1}^{k} X_{n_j+n}^{-1}(B_j)\right).$$

Un processo stazionario corrisponde ad una trasformazione che conserva la misura nel seguente modo. Si considerino lo spazio delle successioni reali doppiamente infinite.

$$\mathbf{R}^{\mathbf{Z}} := \{(\dots, x_{-1}, x_0, x_1, \dots, x_n, \dots) : x_n \in \mathbf{R}\ (n \in \mathbf{Z})\}$$

e la trasformazione $\varphi : \Omega \to \mathbf{R}^{\mathbf{R}}$ definita, per ogni $n \in \mathbf{Z}$, da

$$(\varphi\omega)_n := X_n(\omega).$$

Si definisca ora una misura di probabilità sui boreliani di $\mathbf{R}^{\mathbf{Z}}$ mediante

$$\nu(B) := \mu\left(\varphi^{-1}(B)\right).$$

Infine, sia $T : \mathbf{R}^{\mathbf{Z}} \to \mathbf{R}^{\mathbf{Z}}$ la trasformazione bilatera

$$(Tx)_n := x_{n+1} \qquad (n \in \mathbf{Z}).$$

Come conseguenza della stazionarietà del processo $\{X_n : n \in \mathbf{Z}\}$ la probabilità ν è invariante rispetto a T e, quindi, su tutto $\mathcal{B}\left(|R^{\mathbf{Z}}\right)$; considerata la proiezione canonica $\pi_n : \mathbf{R}^{\mathbf{Z}} \to \mathbf{R}$,

$$\pi_n\{x_k : k \in \mathbf{Z}\} := x_n \qquad (n \in \mathbf{Z}).$$

Allora $\{\pi_n\}$ ha su $\mathbf{R}^{\mathbf{Z}}$ le stesse leggi congiunte di $\{X_n\}$ su Ω. Pertanto ogni processo stocastico stazionario deriva da una trasformazione che conserva la misura su $\mathbf{R}^{\mathbf{Z}}$.

4. La ricorrenza

La teoria ergodica si interessa delle proprietà asintotiche delle trasformazioni che conservano la misura, vale a dire delle proprietà della successione $\{T^n\}$. Perciò si studiano necessariamente trasformazioni di uno spazio in sé, $T : \Omega \to \Omega$.

Il primo risultato fu dato da POINCARÉ in termini di probabilità introducendo, implicitamente la nozione di additività numerabilie.

DEFINIZIONE 4.1. Siano T una trasformazione su $(\Omega, \mathcal{F}, \mu)$ che conserva la misura, e A un insieme di $\mathcal{F}$; un punto ω di A si dice *ricorrente*, rispetto a T e A, se $T^i\,\omega$ appartiene ad A per almeno un numero naturale i; si dice, invece, che ω è *fortemente ricorrente* se $T^n\,\omega$ appartiene ad A per infiniti naturali, cioè $\omega \in \limsup_{n\to+\infty} T^{-n} A$.

TEOREMA 4.1 (di ricorrenza di POINCARÉ). *Siano* $(\Omega, \mathcal{F}, \mu)$ *uno spazio di probabilità,* $T : \Omega \to \Omega$ *una trasformazione che conserva la misura e* A *un insieme di* $\mathcal{F}$. *Allora, quasi tutti i punti di* A *sono ricorrenti.*

DIMOSTRAZIONE. Se $\mu(A) = 0$ non vi è nulla da dimostrare. Si supponga, allora, che sia $\mu(A) > 0$. Si indichi con B il sottoinsieme di A composto dai punti di A che non tornano mai in A. Si vuole dimostrare che $\mu(B) = 0$. Ora

$$B = A \setminus \cup_{n\in\mathbf{N}} \{\omega \in A : T^n \omega \in A\} = A \setminus \cup_{n\in\mathbf{N}} T^{-n} A = A \cap \left(\cap_{n\in\mathbf{N}} T^{-n} A^c\right),$$

ciò che prova, intanto che B è misurabile, $B \in \mathcal{F}$. Se $0 \le i < j$ si ha

$$T^{-j} B \cap T^{-i} B \subset T^{-j} A \setminus \cup_{n\in\mathbf{N}} T^{-i-n} A = \emptyset.$$

Ne segue che la successione $\{T^{-n} B : n \in \mathbf{Z}_+\}$ è composta da insieme misurabili a due a due disgiunti. Ma $\mu\left(T^{-n} B\right) = \mu(B)$, per ogni $n \in \mathbf{Z}_+$, perché T conserva la misura. Da

$$\Omega \supset \cup_{n\in\mathbf{Z}_+} T^{-n} B$$

scende che

$$1 = \mu(\Omega) \ge \mu\left(\cup_{n\in\mathbf{Z}_+} T^{-n} B\right) = \sum_{n\in\mathbf{Z}_+} \mu\left(T^{-n} B\right) = \sum_{n\in\mathbf{Z}_+} \mu(B);$$

ma ciò è possibile solo se $\mu(B) = 0$. □

TEOREMA 4.2 (di ricorrenza forte). *Nelle stesse ipotesi del teorema precedente si supponga che* A *sia un insieme misurabile con* $\mu(A) > 0$. *Allora quasi tutti i punti di* A *sono fortemente ricorrenti.*

DIMOSTRAZIONE. L'insieme dei punti che entrano infinite volte in A è dato da

$$B := \limsup_{n\to+\infty} T^{-n} A,$$

mentre l'insieme dei punti di A che sono fortemente ricorrenti è $A \cap B$. Posto, per $n \in \mathbf{Z}_+$,

$$C_n := \cup_{k\ge n} T^{-k} A,$$

si ha $T^{-1} C_n = C_{n+1}$ e

$$C_0 \supset C_1 \supset \cdots \supset C_n \supset C_{n+1} \supset \ldots$$

con $C_0 \supset A$. Poiché T conserva la misura si ha, per ogni $n \in \mathbf{N}$,

$$\mu\left(C_{n+1}\right) = \mu\left(T^{-1} C_n\right) = \mu\left(C_n\right) = \cdots = \mu\left(C_0\right).$$

Ora

$$\mu(B) = \mu\left(\cap_{n\in\mathbf{Z}_+} C_n\right) = \lim_{n\to+\infty} \mu\left(C_n\right) = \mu\left(C_0\right). \tag{4.1}$$

D'altro canto

$$A = A \cap C_0 = (A \cap B) \cup \left[A \cap \left(C_0 \setminus B\right)\right],$$

che è un'unione disgiunta, sicché

$$\mu(A \cap B) = \mu(A \cap C_0) - \mu[A \cap (C_0 \setminus B)].$$

Poiché B è incluso in C_0, $B \subset C_0$, in virtú della (4.1) si ha

$$\mu[A \cap (C_0 \setminus B)] \leq \mu(C_0 \setminus B) = \mu(C_0) - \mu(B) = 0;$$

pertanto

$$\mu(A \cap B) = \mu(A \cap C_0) = \mu(A),$$

ciò che conclude la dimostrazione. □

L'ipotesi che la trasformazione T conservi la misura è stata usata solo per stabilire che la successione $\{T^{-n} B : n \in \mathbf{N}\}$ è costituita da insiemi a due a due disgiunti e per escludere la possibilità che B abbia misura diversa da zero. Pertanto tutti i teoremi dati possono essere rafforzati eliminando le ipotesi che non sono essenziali.

Definizione 4.2. Data una trasformazione $T : \Omega \to \Omega$, si dice che un insieme $A \subset \Omega$ è *vagante* (in inglese *wandering*) se la successione $\{T^{-n} A : n \in \mathbf{Z}_+\}$ è disgiunta.

Definizione 4.3. Una trasformazione misurabile T su $(\Omega, \mathcal{F}, \mu)$ si dice *dissipativa* se esiste un insieme misurabile vagante di misura strettamente positiva; altrimenti si dice *conservativa.*

La dimostrazione del teorema di ricorrenza mostra che se T conserva la misura, essa è anche conservativa.

Definizione 4.4. Una trasformazione misurabile T su $(\Omega, \mathcal{F}, \mu)$ si dice *comprimibile*, se esiste un insieme $A \in \mathcal{F}$ tale che

$$A \subset T^{-1} A \qquad \text{e} \qquad \mu\left(T^{-1} A \setminus A\right) > 0;$$

se un tale insieme non esiste T si dice *incomprimibile.*

Teorema 4.3. *Per una trasformazione misurabile T su $(\Omega, \mathcal{F}, \mu)$ sono equivalenti le affermazioni*

(a) *T è comprimibile;*
(b) *T è dissipativa.*

Dimostrazione. (a) $\Longrightarrow$ (b) Si supponga che T sia comprimibile, sia A un insieme come nella definizione 4.4; posto $B := T^{-1} A \setminus A$, risulta $\mu(B) > 0$.

Inoltre, per ogni $n \in \mathbf{N}$, si ha

$$\begin{aligned} B \cap T^{-n} B &= \left(T^{-1} A \setminus A\right) \cap \left(T^{-n-1} A \setminus T^{-n} A\right) \\ &= T^{-1} A \cap A^c \cap T^{-n-1} A \cap T^{-n} A^c \\ &= T^{-1} \left(A \cap T^{-n} A\right) \bigcap \left(A^c \cap T^{-n} A^c\right) = T^{-1} A \cap T^{-n} A^c \\ &= T^{-1} \left(A \cap T^{-n+1} A^c\right) = T^{-1} \emptyset = \emptyset, \end{aligned}$$

perché da $A \subset T^{-1} A$ scendono le inclusioni $T^{-1} A^c \subset A^c$, $T^{-n} A \supset A$ e $T^{-n} A^c \subset A^c$. Ma allora la successione $\{T^{-n} B : n \in \mathbf{Z}_+\}$ è disgiunta e quindi T è dissipativa.

(b) $\Longrightarrow$ (a) Sia T dissipativa; esiste pertanto un insieme $B \in \mathcal{F}$ tale che sia $\mu(B) > 0$ e $B \cap T^{-n} B = \emptyset$ per ogni $n \in \mathbf{N}$. Posto

$$A := \left(\cup_{n \in \mathbf{Z}_+} T^{-n} B\right)^c = \cap_{n \in \mathbf{Z}_+} T^{-n} B^c,$$

si ha

$$T^{-1} A = \cap_{n \in \mathbf{N}} T^{-n} B^c \supset \cap_{n \in \mathbf{Z}_+} T^{-n} B^c = A.$$

Inoltre

$$\begin{aligned} T^{-1} A \setminus A &= \cap_{n \in \mathbf{N}} T^{-n} B^c \setminus \cap_{n \in \mathbf{Z}_+} T^{-n} B^c \\ &= \left(\cup_{n \in \mathbf{N}} T^{-n} B\right)^c \setminus \left(\cup_{n \in \mathbf{Z}_+} T^{-n} B\right)^c \\ &= \left(\cup_{n \in \mathbf{N}} T^{-n} B\right)^c \cap \left(\cup_{n \in \mathbf{Z}_+} T^{-n} B\right) = B, \end{aligned}$$

cosicché $\mu\left(T^{-1} A \setminus A\right) = \mu(B) > 0$ e T risulta essere comprimibile. □

5. Ergodicità

Diamo qui la definizione di trasformazione ergodica.

Definizione 5.1. Dato uno spazio mensurale $(\Omega, \mathcal{F}, \mu)$, una trasformazione T di Ω in sé che conservi la misura si dice *ergodica* se ogni insieme misurabile che sia invariante rispetto a T, cioè tale che $A \in \mathcal{F}$ e $A = T^{-1} A$, ha misura nulla oppure il suo complementare ha misura nulla, $\mu(A) = 0$ oppure $\mu\left(A^c\right) = 0$.

Osservazione 5.1. Se μ è una misura di probabilità, T è ergodica se si ha $\mu(A) = 0$ oppure $\mu(A) = 1$ per ogni insieme misurabile invariante A.

Il seguente teorema dà formulazioni equivalenti dell'ergodicità.

Teorema 5.1. *Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio di probabilità e $T : \Omega \to \Omega$ una trasformazione che conserva la misura. Le seguenti affermazioni sono equivalenti:*

(a) *T è ergodica;*
(b) *se B è un insieme misurabile tale che $\mu\left(T^{-1} B \Delta B\right)=0$, allora si ha $\mu(B)=0$ oppure $\mu(B)=1$;*
(c) *se gli insiemi misurabili A e B hanno entrambi probabilità strettamente positiva, $\mu(A)>0$ e $\mu(B)>0$, allora esiste un numero naturale n tale che $\mu\left(T^{-n} A \cap B\right)>0$;*
(d) *se $A \in \mathcal{F}$ è tale che $\mu(A)>0$ e se $\widetilde{A}:=\cup_{n \in \mathbf{N}} T^{-n} A$, allora $\mu(\widetilde{A})=1$.*

Dimostrazione. (a) $\Longrightarrow$ (b) Sia $B \in \mathcal{F}$ un insieme tale che $\mu\left(T^{-1} B \Delta B\right)=0$ e si ponga

$$C:=\limsup_{n \rightarrow+\infty} T^{-n} B=\cap_{n \in \mathbf{Z}_{+}} \cup_{k \geq n} T^{-k} B.$$

Ovviamente l'insieme C cosí definito è misurabile, $C \in \mathcal{F}$; inoltre esso è invariante, $T^{-1} C=C$, come subito si vede. Ma allora la condizione (a) implica che sia $\mu(C)=0$ oppure $\mu(C)=1$. Basta ora mostrare che è $\mu(C)=\mu(B)$.

Posto $B_{n}:=\cup_{k \geq n} T^{-k} B$, si ha $C=\lim_{n \rightarrow+\infty} B_{n}$ e, per ogni $n \in \mathbf{Z}_{+}$,

$$T^{-1} B_{n}=B_{n+1}, \qquad B_{n} \subset B_{n-1}, \qquad \mu\left(B_{n}\right)=\mu\left(B_{0}\right).$$

Pertanto

$$\cap_{k=0}^{n} B_{k}=B_{n} \quad \text{ e } \quad \mu\left(\cap_{k=0}^{n} B_{k}\right)=\mu\left(B_{n}\right)=\mu\left(B_{0}\right),$$

onde anche $\mu(C)=\mu\left(B_{0}\right)$. D'altro canto, scrivendo B_{0} come un'unione disgiunta si ottiene

$$B_{0}=\cup_{n \in \mathbf{Z}_{+}} T^{-n} B=B \cup\left(T^{-1} B \setminus B\right) \cup\left[T^{-2} B \setminus T^{-1} B \setminus B\right] \cup \ldots$$

e di qui, ricorrendo alla condizione (b),

$$\begin{aligned}
\mu\left(B_{0}\right) & =\mu(B)+\sum_{n \in \mathbf{N}} \mu\left(T^{-n} B \setminus T^{-(n-1)} B \setminus \ldots \setminus B\right) \\
& \leq \mu(B)+\sum_{n \in \mathbf{N}} \mu\left(T^{-n} B \setminus T^{-(n-1)} B\right) \\
& =\mu(B)+\sum_{n \in \mathbf{N}} \mu\left[T^{-(n-1)}\left(T^{-1} B \setminus B\right)\right] \\
& =\mu(B)+\sum_{n \in \mathbf{N}} \mu\left(T^{-1} B \setminus B\right)=\mu(B).
\end{aligned}$$

quindi è anche $\mu(C)=\mu(B)$.

(b) $\Longrightarrow$ (c) Si supponga che sia $\mu(A)\, \mu(B)>0$ e, al tempo stesso, che la (c) sia falsa, vale a dire che, per ogni $n \in \mathbf{N}$, sia

$$\mu\left(T^{-n} A \cap B\right)=0.$$

Ne segue che $\mu\left[(\cup_{n\in\mathbf{N}}T^{-n}\,A)\cap B\right]=0$, perché

$$0\leq\mu\left[\left(\cup_{n\in\mathbf{N}}T^{-n}\,A\right)\cap B\right]=\mu\left[\cup_{n\in\mathbf{N}}\left(T^{-n}\,A\cap B\right)\right]$$
$$\leq\sum_{n\in\mathbf{N}}\mu\left(T^{-n}\,A\cap B\right)=0.$$

Ponendo

$$A_1:=\cup_{n\in\mathbf{N}}T^{-n}\,A,$$

si ha

$$T^{-1}\,A_1\subset A_1\quad\text{e}\quad\mu\left(T^{-1}\,A_1\right)=\mu\left(A_1\right),$$

di modo che

$$\mu\left(T^{-1}\,A_1\Delta A_1\right)=\mu\left(A_1\setminus T^{-1}\,A_1\right)=0.$$

Quindi la condizione (b) implica $\mu(A_1)=0$ oppure $\mu(A_1)=1$. Però $T^{-1}\,A\subset A_1$, onde

$$0<\mu(A)=\mu\left(T^{-1}\,A\right)\leq\mu\left(A_1\right),$$

che dà $\mu(A_1)=1$. Ora le tre relazioni

$$\mu(B)>0,\qquad\mu\left(A_1\right)=1,\qquad\text{e}\qquad\mu\left(A_1\cap B\right)=\mu\left[\left(\cup_{n\in\mathbf{N}}T^{-n}\,A\right)\cap B\right]=0,$$

delle quali la prima vale per ipotesi, mentre le rimanenti due sono appena state dimostrate, non sono compatibili. Infatti, dall'ultima scende

$$1=\mu\left[(A_1\cap B)^c\right]=\mu\left(A_1^c\cup B^c\right)\leq\mu\left(A_1^c\right)+\mu\left(B^c\right)=\mu\left(B^c\right)=1-\mu(B),$$

cioè $\mu(B)=0$, una contraddizione.

(c) $\Longrightarrow$ (a) Si supponga, se possibile, che la (a) sia falsa; esiste, pertanto, un insieme misurabile e invariante A, $A\in\mathcal{F}$ e $T^{-1}\,A=A$ con $\mu(A)\in\,]0,1[$. Si scelga $B=A^c$, di modo che $\mu(B)=1-\mu(A)>0$. Allora per ogni $n\in\mathbf{N}$ si avrebbe

$$\mu\left(T^{-n}\,A\cap A^c\right)=\mu\left(A\cap A^c\right)=\mu(\emptyset)=0,$$

che contraddice alla (c).

(b) $\Longrightarrow$ (d) L'insieme misurabile $\widetilde{A}$ è, ovviamente, misurabile; inoltre si ha

$$T^{-1}\widetilde{A}=\cup_{n=2}^{\infty}T^{-n}A\subset\cup_{n\in\mathbf{N}}T^{-n}A=\widetilde{A}.$$

D'altra parte, poiché T conserva la misura, si ha

$$\mu(\widetilde{A}\Delta T^{-1}\widetilde{A})=\mu(\widetilde{A})-\mu(T^{-1}\widetilde{A})=0,$$

sicché $\mu(\widetilde{A})=1$.

(d) $\Longrightarrow$ (c) Siano A e B insiemi misurabili con $\mu(A)>0$ e $\mu(B)>0$; si supponga, se possibile, che sia $\mu(T^{-1}A\cap B)=0$ per ogni $n\in\mathbf{N}$. Allora

$$\mu(\widetilde{A}\cap B)=\sum_{n\in\mathbf{N}}\mu\left(T^{-n}A\cap B\right)=0.$$

Ma $\widetilde{A} = (\widetilde{A} \cap B) \cup (\widetilde{A} \cap B^c)$; per quanto appena provato si ha

$$1 = \mu(\widetilde{A}) = \mu(\widetilde{A} \cap B^c) \leq \mu\,(B^c)\,,$$

onde $\mu\,(B^c) = 1$ e $\mu(B) = 0$, una contraddizione. □

Se la trasformazione T è anche invertibile il teorema precedente può essere modificato come segue. La dimostrazione è una lieve variazione di quella appena data.

TEOREMA 5.2. *Se T è anche invertibile le seguenti affermazioni sono equivalenti tra loro e sono equivalenti alle condizioni (a) e (b) del teorema precedente:*

(c') *se gli insiemi misurabili A e B hanno entrambi probabilità strettamente positiva, $\mu(A) > 0$ e $\mu(B) > 0$, allora esiste un numero intero $n \in \mathbf{Z}$ tale che $\mu\,(T^{-n}\,A \cap B) > 0$;*
(d) *se $A \in \mathcal{F}$ è tale che $\mu(A) > 0$ e se $\widetilde{A} := \cup_{n \in \mathbf{Z}} T^{-n} A$, allora $\mu(\widetilde{A}) = 1$.*

L'ergodicità può essere caratterizzata usando metodi funzionali che sono spesso utili.

TEOREMA 5.3. *Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio di probabilità e $T : \Omega \to \Omega$ una trasformazione che conserva la misura. Le seguenti affermazioni sono equivalenti:*

(a) *T è ergodica;*
(b) *se $f : \Omega \to \mathbf{C}$ è una funzione misurabile e invariante rispetto a T, vale a dire tale che $(f \circ T)\,(\omega) = f(\omega)$ per ogni $\omega \in \Omega$, allora f è costante q.c.;*
(c) *se $f : \Omega \to \mathbf{C}$ è una funzione misurabile e tale che $(f \circ T)\,(\omega) = f(\omega)$ per quasi ogni $\omega \in \Omega$, allora f è costante q.c.;*
(d) *se $f \in L^2$ è invariante, $(f \circ T)\,(\omega) = f(\omega)$ per ogni $\omega \in \Omega$, allora f è costante q.c.;*
(e) *se $f \in L^2$ è tale che $(f \circ T)\,(\omega) = f(\omega)$ per quasi ogni $\omega \in \Omega$, allora f è costante q.c..*

DIMOSTRAZIONE. Sono ovviamente vere le implicazioni (c) $\Longrightarrow$ (b) $\Longrightarrow$ (d) e (c) $\Longrightarrow$ (e) $\Longrightarrow$ (d).

Per completare la dimostrazione del teorema basterà, quindi, stabilire le implicazioni (a) $\Longrightarrow$ (c) e (d) $\Longrightarrow$ (a).

(a) $\Longrightarrow$ (c) Non è restrittivo supporre che la funzione f sia a valori reali, perché, altrimenti, si possono considerare separatamente le parti reali e immaginaria di f. Per ogni $n \in \mathbf{N}$ e per ogni $k \in \mathbf{Z}$, si definiscano gli insiemi

$$A(k,n) := \left\{\omega \in \Omega : \frac{k}{2^n} \leq f(\omega) < \frac{k+1}{2^n}\right\}.$$

Vale l'inclusione

$$T^{-1}\,A(k,n)\Delta A(k,n) \subset \{f \circ T \neq f\}$$

onde $\mu\left[T^{-1}\,A(k,n)\Delta A(k,n)\right] = 0$. Poiché T è ergodica, il Teorema5.1 assicura che, per ogni $n \in \mathbf{N}$ e per ogni $k \in \mathbf{Z}$, sia $\mu\left(A(k,n)\right) = 0$ oppure $\mu\left(A(k,n)\right) = 1$. Per ogni $n \in \mathbf{N}$ si ha

$$\cup_{k \in \mathbf{Z}} A(k,n) = \Omega,$$

che è un'unione disgiunta. Perciò, per ogni $n \in \mathbf{N}$, esiste un unico indice $k_n \in \mathbf{Z}$ tale che $\mu\left(A(k_n,n)\right) = 1$, mentre $\mu\left(A(k,n)\right) = 0$, se $k \neq k_n$. Si ponga

$$B := \cap_{n \in \mathbf{N}} A\left(k_n, n\right).$$

Evidentemente è $\mu(B) = 1$; inoltre f è costante in B. Infatti, siano ω e ω' due punti di B; allora, per ogni $n \in \mathbf{N}$, è

$$0 \leq f(\omega) - f(\omega') \leq \frac{1}{2^n};$$

l'arbitrarietà di n dà ora $f(\omega) = f(\omega')$.

(d) $\Longrightarrow$ (a) Si considerino l'insieme invariante A, $T^{-1}\,A = A$ e la sua funzione indicatrice 1_A che, ovviamente, appartiene a L^2. Si ha, per ogni $\omega \in \Omega$,

$$1_A(\omega) = 1_{T^{-1}\,A}(\omega) = 1_A(T\,\omega) = \left(1_A \circ T\right)(\omega),$$

cosicché 1_A è q.c. eguale a 0 oppure a 1; integrando, si ha cosí $\mu(A) = 0$ oppure $\mu(A) = 1$. $\square$

6. Esempî di trasformazioni ergodiche

Ritorniamo in questa sezione agli esempî della Sezione 2 per esaminare quali delle trasformazioni che conservano la misura lí presentate siano ergodiche.

Esempio 6.1. L'identità I sullo spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ è ergodica se, e solo se, ogni insieme di $\mathcal{F}$ ha probabilità eguale a 0 o a 1.

Per quanto riguarda le rotazioni della circonferenza unitaria in $\mathbf{C}$, Esempio 3.2, vale il seguente

TEOREMA 6.1. *Per la rotazione* $T_a z = a z$ $(a \in \mathcal{K})$ *sono equivalenti le affermazioni:*

(a) T_a *è ergodica;*
(b) a *non è una radice dell'unità.*

DIMOSTRAZIONE. (a) $\Longrightarrow$ (b) Si supponga che a sia una radice dell'unità; esiste, perciò $p \in \mathbf{N}$ tale che $a^p = 1$; si consideri ora la funzione $f : \mathcal{K} \longrightarrow \mathcal{K}$ definita da $f(z) := z^p$; questa è ovviamente invariante rispetto a T_a, $(f \circ T_a)(z) = f(z)$, ma non è costante sicché T_a non può essere ergodica.

(b) $\Longrightarrow$ (a) Si supponga che a non sia una radice dell'unità e che f sia invariante rispetto a T_a, $f \circ T_a = f$. Si scriva la serie di FOURIER di f

$$f(z) = \sum_{n \in \mathbf{Z}} c_n z^n.$$

Allora

$$(f \circ T_a)(z) = f(a z) = \sum_{n \in \mathbf{Z}} c_n a^n z^n$$

sicché l'invarianza di f dà la relazione, valida per ogni $n \in \mathbf{Z}$,

$$c_n (a^n - 1) = 0,$$

che, per $n \neq 0$, dà, a sua volta $c_n = 0$. Dunque, $f(z) = c_0$, vale a dire che f è costante e, quindi, T_a è ergodica. □

ESEMPIO 6.2. La traslazione bilatera (Esempio 3.6) è ergodica. Sia $\mathcal{A}$ l'algebra di tutte le unioni finite di rettangoli misurabili e sia B un insieme misurabile invariante, $T^{-1} B = B$, $B \in \mathcal{F}$. Sia dato $\varepsilon > 0$ e si scelga $A \in \mathcal{A}$ in modo che sia $\mu(B \Delta A) < \varepsilon$. Si tenga presente che quali che siano gli insiemi misurabili E e F vale la relazione

$$\begin{aligned}|\mu(E) - \mu(F)| &= |\mu(E \cap F) + \mu(E \setminus F) - \mu(F \cap E) - \mu(F \setminus E)| \\ &\leq \mu(E \setminus F) + \mu(F \setminus E) = \mu(E \Delta F).\end{aligned}$$

Ponendo $E = A$ e $F = B$ si ottiene

$$|\mu(B) - \mu(A)| = \mu(B \Delta A) < \varepsilon.$$

Si può scegliere un numero naturale n_0 tale che l'insieme $C = T^{-n_0} A$ dipenda da coordinate distinte da quelle di A. Pertanto

$$\mu(C \cap A) = \mu(C)\, \mu(A) = \mu^2(A).$$

Inoltre

$$\mu(B \Delta C) = \mu\left(T^{-n_0} B \Delta T^{-n_0} A\right) = \mu(B \Delta A) < \varepsilon$$

e, poiché $B \Delta (A \cap C) \subset (B \Delta A) \cup (B \Delta C)$, si ha $\mu\left(B \Delta (A \cap C)\right) < 2\,\varepsilon$. Di qui

$$|\mu(B) - \mu(A \cap C)| < 2\,\varepsilon$$

e

$$\begin{aligned}\left|\mu(B)-\mu^2(B)\right| &\le |\mu(B)-\mu(A\cap C)|+\left|\mu(A\cap C)-\mu^2(B)\right| \\ &< 2\,\varepsilon+\left|\mu^2(A)-\mu^2(B)\right| \\ &\le 2\,\varepsilon+(\mu(A)+\mu(B))\,\left|\mu(A)-\mu(B)\right| < 4\,\varepsilon.\end{aligned}$$

Siccome $\varepsilon > 0$ è arbitrario, si ha $\mu(B) = \mu^2(B)$, vale a dire $\mu(B) = 0$ oppure $\mu(B) = 1$.

Un ragionamento simile mostra che è ergodica anche la traslazione unilatera (Esempio 3.7)

Nella sezione 9 studieremo quando le traslazioni di MARKOV bilatere (Esempio 3.8) siano ergodiche.

7. Il teorema ergodico di Birkhoff

Si è visto studiando la ricorrenza che quasi ogni punto di un insieme misurabile A torna infinite volte in A stesso. Ha grande significato nella Meccanica Statistica porsi il problema di stabilire quanto "tempo" un punto di A passi nello stesso insieme. Se 1_A è la funzione indicatrice di A, la risposta è data dal limite

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1} 1_A\left(T^j\,\omega\right),$$

se tale limite esiste. A questo problema rispose BIRKHOFF nel 1931 con il seguente teorema del quale riporto la dimostrazione di RIESZ.

TEOREMA 7.1 (Teorema ergodico di BIRKHOFF). *Sia $(\Omega,\mathcal{F},\mu)$ uno spazio mensurale σ-finito, sia $T:\Omega\to\Omega$ una trasformazione che conserva la misura e sia f una funzione di L^1. Allora*

(a) $\frac{1}{n}\sum_{j=0}^{n-1} f\left(T^j\,\omega\right)$ *tende q.o. ad una funzione f^* di L^1;*
(b) *la funzione limite f^* è invariante rispetto a T, vale a dire, $f^*\circ T = f^*$.*

Inoltre se lo spazio $(\Omega,\mathcal{F},\mu)$ è finito, $\mu(\Omega) < +\infty$, si ha

$$\int_\Omega f^*\,d\mu = \int_\Omega f\,d\mu. \tag{7.1}$$

La dimostrazione richiede qualche concetto che sarà utile anche per il seguito.

DEFINIZIONE 7.1. Sia $T : \Omega \to \Omega$ una trasformazione che conserva la misura; sullo spazio delle funzioni f definite in Ω e a valori complessi è definito un operatore U_T mediante

$$\forall \omega \in \Omega \qquad (U_T\, f)\,(\omega) := (f \circ T)(\omega) = f(T\,\omega).$$

L'operatore U_T si dice *indotto* dalla trasformazione T.

OSSERVAZIONE 7.1. Se f appartiene a L^p, con $p \geq 1$, allora vi appartiene anche $U_T\, f$, cioè $U_T\, L^p \subset L^p$. Infatti

$$\|U_T\, f\|_p^p = \int_\Omega |f \circ T|^p \, d\mu = \int_\Omega |f|^p \, d\left(\mu\, T^{-1}\right) = \int_\Omega |f|^p \, d\mu = \|f\|_p^p.$$

Da quest'ultima relazione scende che U_T è un'isometria su L^p. In particolare U_T è un'isometria sullo spazio di HILBERT L^2. Non è inutile ricordare che questo implica inoltre che sia

$$\langle U_T f, U_T g\rangle = \langle f, g\rangle$$

per ogni coppia di elementi f e g di L^2.

TEOREMA 7.2 (Teorema ergodico massimale). *Sia $U : L^1 \to L^1$ un operatore lineare positivo, $U\, f \geq 0$ se $f \geq 0$, e contrattivo, $\|U\| \leq 1$ e sia N un numero naturale. Posto*

$$f_0 := 0, \qquad f_n := \sum_{j=0}^{n-1} U^j\, f, \qquad e \qquad F_N := \max\{f_n : 0 \leq n \leq N\}.$$

Allora

$$\int_{\{F_N > 0\}} f\, d\mu \geq 0. \tag{7.2}$$

e

$$\int_{\cup_{N\in\mathbf{N}}\{F_N>0\}} f\, d\mu \geq 0. \tag{7.3}$$

DIMOSTRAZIONE. Poiché, per ogni n con $0 \leq n \leq N$, si ha $F_N \geq f_n$ anche $U\, F_N \geq U\, f_n$. Ora

$$U\, f_n = U\left(\sum_{j=0}^{n-1} U^j\, f\right) = \sum_{j=0}^{n-1} U^{j+1}\, f = f_{n+1} - f \qquad (0 \leq n \leq N),$$

cosicché

$$U\, F_N + f \geq U\, f_n + f = f_{n+1}.$$

Perciò, se $F_N(\omega) > 0$, è

$$\begin{aligned} U\,F_N(\omega) + f(\omega) &\geq \max\{f_{n+1}(\omega) : 0 \leq n \leq N-1\} \\ &= \max\{f_n(\omega) : 1 \leq n \leq N\} = F_N(\omega). \end{aligned}$$

Pertanto nell'insieme $A_N := \{F_N > 0\}$ si ha $f \geq F_N - U\,F_N$ e, quindi, dato che $F_N = 0$ in A_N^c,

$$\begin{aligned} \int_{A_N} f\,d\mu &\geq \int_{A_N} F_N\,d\mu - \int_{A_N} U\,F_N\,d\mu = \int_{\Omega} F_N\,d\mu - \int_{A_N} U\,F_N\,d\mu \\ &\geq \int_{\Omega} F_N\,d\mu - \int_{\Omega} U\,F_N\,d\mu \\ &= \|F_N\|_1 - \|U\,F_N\|_1 \geq \|F_N\|_1 - \|F_N\|_1 = 0. \end{aligned}$$

Per dimostrare la (7.3), si scriva

$$\int_{\cup_{n\in\mathbf{N}}\{F_N>0\}} f\,d\mu = \int_{\cup_{n\in\mathbf{N}}\{F_N>0\}} f^+\,d\mu - \int_{\cup_{n\in\mathbf{N}}\{F_N>0\}} f^-\,d\mu$$

e si usi il teorema di convergenza monotona a ciascuno degli ultimi due integrali. □

Si noti che il teorema ergodico massimale può essere applicato, in particolare, all'operatore U_T indotto da una trasformazione T che conserva la misura. Per un tale operatore si ha, infatti, $\|U_T\| \leq 1$.

COROLLARIO 7.1. *Se $T : \Omega \to \Omega$ conserva la misura, se φ è integrabile, $\varphi \in L^1$, se si pone*

$$B_a := \left\{ \sup_{n\in\mathbf{N}} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(T^j\,\omega\right) > a \right\},$$

e se A è un insieme misurabile invariante e di misura strettamente positiva e finita, $A \in \mathcal{F}$, $T^{-1}\,A = A$, $0 < \mu(A) < +\infty$, allora

$$\int_{B_a\cap A} \varphi\,d\mu \geq a\,\mu\,(B_a \cap A)\,. \tag{7.4}$$

DIMOSTRAZIONE. Si supponga dapprima che sia $\mu(\Omega) < +\infty$ e che A coincida con tutto lo spazio Ω. Si ponga $f := \varphi - \alpha$, sicché

$$B_a = \cup_{N\in\mathbf{N}}\,\{F_N > 0\} = \cup_{N\in\mathbf{N}} A_N.$$

La (7.3) dà, per ogni $N \in \mathbf{N}$,

$$\int_{B_a} f\,d\mu \geq 0,$$

onde

$$\int_{B_a} \varphi \, d\mu \geq a \, \mu \, (B_a) \, .$$

Il caso generale segue considerando la restrizione di T ad A. □

DIMOSTRAZIONE DEL TEOREMA DI BIRKHOFF. Si definiscano le funzioni

$$f^* := \limsup_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j,$$

$$f_* := \liminf_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j.$$

Mostriamo che f^* e f_* sono invarianti rispetto a T, $f^* \circ T = f^*$, e $f_* \circ T = f_*$. Si ponga

$$\varphi_n := \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j$$

e si noti che

$$\frac{n+1}{n} \, \varphi_{n+1} - \varphi_n \circ T = \frac{f}{n}. \tag{7.5}$$

Da quest'ultima relazione segue che

$$\begin{aligned} f^* &= \limsup_{n \to +\infty} \frac{n+1}{n} \, \varphi_{n+1} = \limsup_{n \to +\infty} \left(\varphi_n \circ T + \frac{f}{n} \right) \\ &\leq \limsup_{n \to +\infty} \varphi_n \circ T + \limsup_{n \to +\infty} \frac{f}{n} = f^* \circ T. \end{aligned}$$

Similmente, sempre utilizzando la (7.5), si ha

$$\begin{aligned} f^* \circ T &= \limsup_{n \to +\infty} \varphi_n \circ T = \limsup_{n \to +\infty} \left(\frac{n+1}{n} \, \varphi_{n+1} - \frac{f}{n} \right) \\ &\leq \limsup_{n \to +\infty} \frac{n+1}{n} \, \varphi_{n+1} - \liminf_{n \to +\infty} \frac{f}{n} = \limsup_{n \to +\infty} \frac{n+1}{n} \, \varphi_{n+1} = f^*. \end{aligned}$$

Le ultime due diseguaglianze, insieme, danno l'invarianza di f^* rispetto a T. In maniera analoga si mostra che è invariante rispetto a T la funzione f_*.

Siano ora α e β due numeri reali con $\alpha < \beta$ e si ponga

$$A_{\alpha,\beta} := \{ f_* < \alpha < \beta < f^* \} \, .$$

L'insieme $A_{\alpha,\beta}$ è misurabile, $A_{\alpha,\beta} \in \mathcal{F}$, e invariante rispetto a T, $T^{-1} \, A_{\alpha,\beta} = A_{\alpha,\beta}$. Se si dimostrasse che $A_{a,b}$ ha misura finita, $\mu \, (A_{a,b}) < +\infty$, allora si potrebbe usare il Corollario 7.1.

Se $\beta > 0$, in virtú dell'ipotesi che la misura μ sia σ–finita, si può scegliere un insieme misurabile B contenuto in $A_{\alpha,\beta}$ e di misura finita, $B \subset A_{\alpha,\beta}$ e $\mu(B) < +\infty$. Cosí la funzione $h := f - \beta \mathbf{1}_B$ è integrabile e dal Teorema ergodico massimale 7.2, con H_N definita in maniera analoga a F_N, segue che, per ogni $N \in \mathbf{N}$, si ha

$$\int_{\cup_{N\in\mathbf{N}}\{H_N>0\}} (f - \beta \mathbf{1}_B)\, d\mu \geq 0.$$

Si osservi che se il punto ω appartiene a $A_{\alpha,\beta}$ si ha

$$\beta < f^*(\omega) = \limsup_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} f\left(T^j \omega\right),$$

cosicché almeno una delle somme

$$\frac{1}{n} \sum_{j=0}^{n-1} f\left(T^j \omega\right)$$

deve essere maggiore di β. Ne segue che almeno una delle somme

$$\frac{1}{n} \sum_{j=0}^{n-1} \left\{ f\left(T^j \omega\right) - \beta \mathbf{1}_B\left(T^j \omega\right) \right\}$$

è positiva. Pertanto

$$B \subset A_{\alpha,\beta} \subset \cup_{N\in\mathbf{N}}\{H_N > 0\}$$

e

$$\begin{aligned} \int_\Omega |f|\, d\mu \geq \int_B |f|\, d\mu \geq \int_B f\, d\mu &= \int_{B\cap\cup_{N\in\mathbf{N}}\{H_N>0\}} f\, d\mu \\ &\geq \beta\, \mu\left(B \cap \cup_{N\in\mathbf{N}}\{H_N > 0\}\right) = \beta\, \mu(B), \end{aligned}$$

vale a dire $\mu(B) \leq \|f\|_1/\beta$. Ogni sottoinsieme B di $A_{\alpha,\beta}$ che sia misurabile e che abbia misura finita ha la misura maggiorata da $\|f\|_1/\beta$. Ma allora la σ–finitezza dello spazio mensurale $(\Omega, \mathcal{F}, \mu)$ implica che anche $A_{\alpha,\beta}$ abbia misura finita, $\mu\left(A_{\alpha,\beta}\right) < +\infty$. Infatti, si può scrivere

$$A_{\alpha,\beta} = \cup_{n\in\mathbf{N}} E_n,$$

ove gli insiemi E_n ($n \in \mathbf{N}$) sono misurabili, e si può supporre, senza perdita di generalità, che essi siano a due a due disgiunti, $E_j \cap E_k$ ($j \neq k$); inoltre, essi hanno tutti misura finita, $\mu\left(E_n\right) < +\infty$, per ogni $n \in \mathbf{N}$. Ora

$$\mu\left(A_{\alpha,\beta}\right) = \mu\left(\cup_{n\in\mathbf{N}} E_n\right) = \sum_{n\in\mathbf{N}} \mu\left(E_n\right);$$

d'altro canto, ogni somma parziale di tale serie,

$$\sum_{j=1}^{n} \mu\left(E_j\right) = \mu\left(\cup_{j=1}^{n} E_j\right)$$

è limitata da $\|f\|_1/\beta$, perché $\cup_{j=1}^{n} E_j$ è un sottoinsieme misurabile di $A_{\alpha,\beta}$ con misura finita; quindi $\mu(A_{\alpha,\beta}) \leq \|f\|_1/\beta$.

Se, invece, $\beta \leq 0$, allora si ha necessariamente $\alpha < 0$; si può ora applicare lo stesso ragionamento a $-f$ e a $-\alpha$ anziché a f e a β. Cosí si giunge nuovamente alla conclusione $\mu(A_{\alpha,\beta}) < +\infty$.

Si ponga ora

$$B_\beta := \left\{ \sup_{n\in\mathbf{N}} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j > \beta \right\}.$$

Poiché $A_{\alpha,\beta} \subset B_\beta$, la (7.4) dà

$$\int_{A_{\alpha,\beta}} f\, d\mu = \int_{A_{\alpha,\beta}\cap B_\beta} f\, d\mu \geq \beta\, \mu(A_{\alpha,\beta} \cap B_\beta) = \beta\, \mu(A_{\alpha,\beta}). \tag{7.6}$$

Sostituendo f, β, e α con $-f$, $-\alpha$ e $-\beta$ rispettivamente si ottiene $(-f)^* = -f_*$, $(-f)_* = -f^*$ e

$$\int_{A_{\alpha,\beta}} f\, d\mu \leq \alpha\, \mu(A_{\alpha,\beta}). \tag{7.7}$$

Sottraendo la (7.6) dalla (7.7) si ottiene $(\alpha - \beta)\, \mu(A_{\alpha,\beta}) \geq 0$, che, poiché è $\alpha < \beta$, implica $\mu(A_{\alpha,\beta}) = 0$. Ora, esprimendo l'insieme $\{f_* < f^*\}$ come un'unione numerabile,

$$\{f_* < f^*\} = \bigcup_{\substack{\alpha,\beta\in\mathbf{Q} \\ \alpha<\beta}} A_{\alpha,\beta},$$

si ottiene $\mu(\{f_* < f^*\}) = 0$, cioè $f^* = f_*$ q.o.. Cosí si è mostrato che

$$\frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j$$

converge q.o. a $f^* = f_*$.

Rimane da mostrare che il limite f^* è integrabile. A tal fine, si usa il lemma di FATOU. Posto

$$g_n := \left| \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j \right|,$$

si ha

$$0 \leq \int_\Omega g_n\, d\mu = \int_\Omega \left| \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j \right| d\mu \leq \frac{1}{n} \sum_{j=0}^{n-1} \int_\Omega \left| f \circ T^j \right| d\mu = \|f\|_1,$$

sicché, per il lemma di FATOU,

$$\int |f_*| \, d\mu = \int \liminf_{n\to+\infty} g_n \, d\mu \leq \liminf_{n\to+\infty} \int_\Omega g_n \, d\mu \leq \|f\|_1 < +\infty.$$

Perciò $|f_*|$ è integrabile.

(c) Per $k \in \mathbf{Z}$ e $n \in \mathbf{N}$ si ponga

$$A(k,n) := \left\{ \frac{k}{n} \leq f^* < \frac{k+1}{n} \right\}$$

Fissato arbitrariamente $\varepsilon > 0$, si ha $A(k,n) \cap B_{k/n-\varepsilon} = A(k,n)$ e, per il Corollario 7.1,

$$\int_{A(k,n)} f \, d\mu = \int_{A(k,n)\cap B_{k/n-\varepsilon}} f \, d\mu \geq \left(\frac{k}{n} - \varepsilon \right) \mu\left(A(k,n)\right),$$

cioè

$$\int_{A(k,n)} f \, d\mu \geq \frac{k}{n} \, \mu\left(A(k,n)\right).$$

Di qui segue

$$\int_{A(k,n)} f^* \, d\mu \geq \frac{k+1}{n} \, \mu\left(A(k,n)\right) \leq \frac{1}{n} \, \mu\left(A(k,n)\right) + \int_{A(k,n)} f \, d\mu.$$

Da quest'ultima diseguaglianza, sommando su k scende, per ogni $n \in \mathbf{N}$,

$$\int_\Omega f^* \, d\mu \leq \frac{1}{n} \, \mu(\Omega) + \int_\Omega f \, d\mu,$$

e, perciò, poiché $\mu(\Omega) < +\infty$,

$$\int_\Omega f^* \, d\mu \leq \int_\Omega f \, d\mu.$$

Lo stesso ragionamento applicato a $-f$ dà

$$\int_\Omega (-f)^* \, d\mu \leq \int_\Omega (-f) \, d\mu \qquad \text{cioè} \qquad -\int_\Omega f_* \, d\mu \leq -\int_\Omega f \, d\mu.$$

Poiché $f^* = f_*$ q.o., si ha

$$\int_\Omega f_* \, d\mu \geq \int_\Omega f \, d\mu,$$

sicché vale la (7.1). □

OSSERVAZIONE 7.2. Si mostra facilmente con un esempio che l'ipotesi $\mu(\Omega) < +\infty$ che lo spazio abbia misura finita è essenziale affinché valga la (7.1). Si consideri lo spazio mensurale $(\mathbf{R}, \mathcal{B}, \lambda)$, ove λ è la misura di LEBESGUE definita sui boreliani $\mathcal{B}$. La trasformazione $T : \mathbf{R} \to \mathbf{R}$ definita da $Tx := x + 1$ ovviamente conserva la misura. Per f, la funzione indicatrice dell'intervallo $[0, 1]$ si ha

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j = 0 = f^*,$$

sicché

$$\int_\Omega f^* \, d\lambda = 0 \qquad \text{mentre} \qquad \int_\Omega f \, d\lambda = 1.$$

8. Corollarî del teorema di Birkhoff

Il teorema di BIRKHOFF assume una formulazione piú forte quando la trasformazione T è ergodica.

COROLLARIO 8.1. *Nelle stesse ipotesi del Teorema di* BIRKHOFF, *se la trasformazione $T : \Omega \to \Omega$ è ergodica, allora la funzione limite f^* è costante q.o.. In particolare*

(a) *se la misura μ è finita, $\mu(\Omega) < +\infty$, è*

$$f^* = \frac{1}{\mu(\Omega)} \int_\Omega f \, d\mu; \tag{8.1}$$

(b) *se la misura μ è infinita, $\mu(\Omega) = +\infty$, allora $f^* = 0$ q.o..*

DIMOSTRAZIONE. Per il teorema di BIRKHOFF la funzione f^* è invariante rispetto a T, ma le uniche funzioni invarianti quando T è ergodica sono costanti q.o.. Se $\mu(\Omega) < +\infty$, vale

$$\int_\Omega f^* \, d\mu = \int_\Omega f \, d\mu,$$

onde la (8.1). Se, invece, $\mu(\Omega) = +\infty$, allora $f^* = 0$ q.c., perché questa è l'unica funzione che sia, allo stesso tempo, costante ed integrabile. □

OSSERVAZIONE 8.1. Nel caso di uno spazio di probabilità è facile dare una forma propbabilistica alla funzione limite f^* del teorema di BIRKHOFF. Si osserva subito che la famiglia $\mathcal{I}$ degli insiemi invarianti rispetto a T,

$$\mathcal{I} := \left\{ A \in \mathcal{F} : T^{-1}A = A \right\},$$

è una tribú. Inoltre se una funzione $\varphi \in L^1$ è invariante, $\varphi \circ T = \phi$ e se A è un insieme invariante, $T^{-1}A = A$, allora è invariante anche $\phi 1_A$; infatti

$$(\varphi\, 1_A) \circ T = (\varphi \circ T)\ (1_A \circ T) = \varphi\, 1_{T^{-1}A} = \varphi\, 1_A.$$

La (7.1), applicata alla funzione $f\, 1_A$ dà

$$\int_A f^*\, d\mu = \int_A f\, d\mu,$$

sicché la funzione limite f^* è la speranza condizionata di f rispetto alla tribú $\mathcal{I}$ degli insiemi invarianti,

$$f^* = E\left(f \mid \mathcal{I}\right).$$

Il corollario che segue è particolarmente importante perché caratterizza le trasformazioni ergodiche in uno spazio di probabilità.

COROLLARIO 8.2. *Sia $(\Omega, \mathcal{F}, \mu)$ uno spazio di probabilità e sia $T : \Omega \to \Omega$ una trasformazione che conserva la misura. Sono allora equivalenti le affermazioni*

(a) *T è ergodica;*
(b) *per ogni coppia A e B di insiemi misurabili, $A, B \in \mathcal{F}$, si ha*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(T^{-j}\, A \cap B\right) = \mu(A)\, \mu(B). \tag{8.2}$$

DIMOSTRAZIONE. (a) $\Longrightarrow$ (b) Sia T ergodica. Se $f = 1_A$ con A misurabile, il teorema di BIRKHOFF e il Corollario 8.1 danno

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} 1_A \circ T^j = \mu(A) \qquad \text{q.c.},$$

relazione dalla quale, moltiplicando ambo i membri per 1_B con $B \in \mathcal{F}$, scende

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left(1_A \circ T^j\right)\, 1_B = \mu(A)\, 1_B \qquad \text{q.c.}.$$

L'asserto segue ora integrando e ricorrendo al teorema di convergenza dominata.

(b) $\Longrightarrow$ (a) Sia A un insieme misurabile invariante rispetto a T, $A \in \mathcal{F}$ e $T^{-1}\, A = A$; prendendo $B = A$ nella (8.2), si ottiene

$$\mu(A) = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(A) = [\mu(A)]^2,$$

donde $\mu(A) = 0$ oppure $\mu(A) = 1$. $\square$

COROLLARIO 8.3 (Teorema ergodico medio di VON NEUMANN). *Sia T una trasformazione che conserva la misura sullo spazio di probabilità $(\Omega, \mathcal{F}, \mu)$. Se f è una funzione di L^p, con $p \in [0, +\infty[$, allora esiste una funzione $f^* \in L^p$, invariante rispetto a T, $U_T f^* = f^* \circ T = f^*$ tale che*

$$\lim_{n\to+\infty} \left\| \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j - f^* \right\|_p = 0. \tag{8.3}$$

DIMOSTRAZIONE. Sia $\varphi : \Omega \to \mathbf{R}$ una funzione misurabile e limitata, sicché φ appartiene a L^∞ e, quindi, a L^p per ogni $p \in [0, +\infty]$. Scende dal teorema di BIRKHOFF che

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ T^j = \varphi^* \qquad \text{q.c.}$$

ove φ^* è una funzione di L^∞ e quindi di L^p per ogni $p \in [0, +\infty]$. Allora

$$\lim_{n\to+\infty} \left| \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ T^j - \varphi^* \right|^p = 0 \qquad \text{q.c.}$$

sicché il teorema di convergenza dominata dà

$$\lim_{n\to+\infty} \left\| \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ T^j - \varphi^* \right\|_p = 0.$$

Fissato arbitrariamente $\varepsilon > 0$, esiste cosí un numero naturale $N(\varepsilon, \varphi)$ tale che per ogni $n \geq N(\varepsilon, \varphi)$ e per ogni $k \in \mathbf{N}$, si abbia

$$\left\| \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ T^j - \frac{1}{n+k} \sum_{j=0}^{n+k-1} \varphi \circ T^j \right\|_p < \frac{\varepsilon}{2}.$$

Per $f \in L^p$ con $p \in [0, +\infty[$ si ponga

$$S(n, f) := \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j.$$

La successione $\{S(n, f) : n \in \mathbf{N}\}$ è di CAUCHY in L^p. Intanto, si ha

$$\|S(n, f)\|_p = \left\| \frac{1}{n} \sum_{j=0}^{n-1} U_T^j f \right\|_p \leq \frac{1}{n} \sum_{j=0}^{n-1} \left\| U_T^j f \right\|_p = \frac{1}{n} \sum_{j=0}^{n-1} \| f \|_p = \| f \|_p$$

Si scelga $\varphi \in L^\infty$ in modo che sia $\|f-\varphi\|_p < \varepsilon/4$. Allora, per ogni $n \geq N(\varepsilon, \varphi)$ e per ogni $k \in \mathbf{N}$, è

$$\begin{aligned} &\|S(n,f) - S(n+k,f)\|_p \\ &\qquad \leq \|S(n,f) - S(n,\varphi)\|_p + \|S(n,\varphi) - S(n+k,\varphi)\|_p \\ &\qquad\qquad + \|S(n+k,\varphi) - S(n+k,f)\|_p \\ &\qquad < \|f-\varphi\|_p + \frac{\varepsilon}{2} + \|f-\varphi\|_p < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Siccome L^p è completo, esiste $f^* \in L^p$ tale che

$$\lim_{n\to+\infty} \left\| \frac{1}{n} \sum_{j=0}^{n-1} U_T^j f - f^* \right\|_p = 0.$$

Infine, poiché

$$\frac{n+1}{n} S(n+1,f) - S(n, f\circ T) = \frac{f}{n},$$

f^* risulta essere invariante, cioè $U_T\, f^* = f^*$. □

9. Proprietà piú forti dell'ergodicità

Si è visto che in uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ una trasformazione $T : \Omega \to \Omega$ che conserva la misura è ergodica se, e solo se, per ogn scelta degli insiemi A e B in $\mathcal{F}$, vale la (8.2).

Definizione 9.1. Dato uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$, una trasformazione $T : \Omega \to \Omega$ che conserva la misura si dice

(a) *totalmente ergodica* se, per ogni $k \in \mathbf{N}$, è ergodica la trasformazione T^k;

(b) *debolmente mescolante* se, per ogn scelta degli insiemi A e B in $\mathcal{F}$, è

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \mu\left(T^{-j} A \cap B\right) - \mu(A)\,\mu(B) \right| = 0; \tag{9.1}$$

(c) *fortemente mescolante* se, per ogn scelta degli insiemi A e B in $\mathcal{F}$, è

$$\lim_{n\to+\infty} \mu\left(T^{-n} A \cap B\right) = \mu(A)\,\mu(B). \tag{9.2}$$

Osservazione 9.1. Nella definizione di mescolanza forte non è, in genere possibile sostituire la misura di probabilità μ con una misura arbitraria. Infatti, se nella relazione

$$\lim_{n\to+\infty} \mu\left(T^{-n} A \cap B\right) = \mu(A)\,\mu(B). \tag{9.3}$$

si prende $A = B = \Omega$ si ottiene $\mu(\Omega) = [\mu(\Omega)]^2$, sicchè, se μ è una misura finita e non banale ($\mu(A) = 0$ per ogni $A \in \mathcal{F}$), è necessariamente $\mu(\Omega) = 1$. Se, invece, μ non è una misura finita, $\mu(\Omega) = +\infty$, si supponga che esista un insieme misurabile A di misura finita e non nulla, $0 < \mu(A) < +\infty$. Allora, prendendo nella (9.3), $B = \Omega$ si trova

$$+\infty > \mu(A) = \mu\left(T^{-n} A\right) = \mu\left(T^{-n} A \cap \Omega\right) \xrightarrow[n \to +\infty]{} \mu(A)\, \mu(\Omega) = +\infty,$$

che è una contraddizione. La (9.3) ha dunque senso, se $\mu(\Omega) = +\infty$, solo per la misura $\mu : \mathcal{F} \to \overline{\mathbf{R}}_+$ definita da

$$\mu(A) := \begin{cases} 0, & A = \emptyset, \\ +\infty, & A \neq \emptyset. \end{cases}$$

Tale misura riveste però scarso interesse.

TEOREMA 9.1. *Siano (Ω, F, μ) uno spazio di probabilità e T una trasformazione che conserva la misura. Allora*

(a) *se T è fortemente mescolante, essa è anche debolmente mescolante;*
(b) *se T è debolmente mescolante, essa è anche totalmente ergodica;*
(c) *se T è totalmente ergodica, essa è anche ergodica.*

DIMOSTRAZIONE. (a) La (9.2) si può scrivere anche nella forma

$$\lim_{n \to +\infty} \left| \mu\left(T^{-n} A \cap B\right) - \mu(A)\, \mu(B) \right| = 0;$$

l'asserto segue immediatamente ricordando che una successione convergente converge anche nel senso di CESÀRO.

(b) Si deve mostrare che, per ogni $k \in \mathbf{N}$ e per ogni scelta di due insiemi misurabili A e B, si ha

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(T^{-k\,j} A \cap B\right) = \mu(A)\, \mu(B).$$

Ora

$$
\begin{aligned}
0 &\leq \left| \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(T^{-kj} A \cap B\right) - \mu(A)\,\mu(B) \right| \\
&= \left| \frac{1}{n} \sum_{j=0}^{n-1} \left[\mu\left(T^{-kj} A \cap B\right) - \mu(A)\,\mu(B)\right] \right| \\
&\leq \frac{1}{n} \sum_{j=0}^{n-1} \left| \mu\left(T^{-kj} A \cap B\right) - \mu(A)\,\mu(B) \right| \\
&\leq \frac{k\,(n-1)}{n} \, \frac{1}{k\,(n-1)} \sum_{i=0}^{k\,(n-1)} \left| \mu\left(T^{-i} A \cap B\right) - \mu(A)\,\mu(B) \right| \xrightarrow[n \to +\infty]{} 0.
\end{aligned}
$$

La (c) è ovvia. □

Rimandiamo, per il momento, la presentazione di esempî che illustrino i varî tipi di trasformazione, anticipando solo, sin da ora, che non sarà facile dare esempî di trasformazioni debolmente mescolanti.

Il teorema che segue mostra che le condizioni di ergodicità, di forte ergodicità, di mescolanza debole e di mescolanza forte possono essere controllate anziché su tutti i sottoinsiemi misurabili solo su quella di una sottoalgebra $\mathcal{A}$ di $\mathcal{F}$ che la generi, $\mathcal{F}(\mathcal{A}) = \mathcal{F}$.

TEOREMA 9.2. *Siano (Ω, F, μ) uno spazio di probabilità, T una trasformazione che conserva la misura e$\mathcal{A}$ un'algebra che genera la tribú $\mathcal{F}$, $\mathcal{F}(\mathcal{A}) = \mathcal{F}$. Allora*

(a) *T è ergodica se, e solo se, la* (8.2) *vale per ogni di insiemi A e B appartenenti all'algebra $\mathcal{A}$;*
(b) *T è debolmente mescolante se, e solo se, la* (9.1) *vale per ogni di insiemi A e B appartenenti all'algebra $\mathcal{A}$;*
(c) *T è fortemente mescolante se, e solo se, la* (9.2) *vale per ogni di insiemi A e B appartenenti all'algebra $\mathcal{A}$.*

DIMOSTRAZIONE. In tutti e tre i casi basta dimostrare che basta controllare le relazioni in questione quando gli insiemi considerati appartengano all'algebra $\mathcal{A}$.

Si osservi che, fissati $\varepsilon > 0$ e gli insiemi A e B in $\mathcal{F}$, è possibile scegliere A_0 e B_0 in $\mathcal{A}$ in maniera che sia

$$\mu\left(A \Delta A_0\right) < \varepsilon \qquad \text{e} \qquad \mu\left(B \Delta B_0\right) < \varepsilon.$$

Pertanto

$$
\begin{aligned}
&\mu\left[\left(T^{-n} A \cap B\right) \Delta\left(T^{-n} A_0 \cap B_0\right)\right] \\
&= \mu\left[\left\{T^{-n} A \Delta\left(T^{-n} A_0 \cap B_0\right)\right\} \cap\left\{B \Delta\left(T^{-n} A_0 \cap B_0\right)\right\}\right] \\
&\leq \mu\left[\left\{T^{-n}\left(A \Delta A_0\right)\right\} \cap\left(T^{-n} A_0 \cap B_0\right)\right]+\mu\left[\left(B \Delta T^{-n} A_0\right) \cap\left(B \Delta B_0\right)\right] \\
&\leq \mu\left[T^{-n}\left(A \Delta A_0\right)\right]+\mu\left(B \Delta B_0\right)=\mu\left(A \Delta A_0\right)+\mu\left(B \Delta B_0\right)<2 \varepsilon .
\end{aligned}
$$

Perciò, posto $C:=T^{-n} A \cap B$, $C_0:=T^{-n} A_0 \cap B_0$ e $D:=C \cap C_0$, si ha

$$
\begin{aligned}
&\left|\mu\left(T^{-n} A \cap B\right)-\mu\left(T^{-n} A_0 \cap B_0\right)\right|=\left|\mu(C)-\mu\left(C_0\right)\right| \\
&\left|\mu(C \setminus D)+\mu(D)-\mu\left(C_0 \setminus D\right)-\mu(D)\right|=\left|\mu(C \setminus D)-\mu\left(C_0 \setminus D\right)\right| \\
&\leq \mu(C \setminus D)+\mu\left(C_0 \setminus D\right)=\mu\left[\left(T^{-n} A \cap B\right) \Delta\left(T^{-n} A_0 \cap B_0\right)\right]<2 \varepsilon .
\end{aligned}
$$

Cosí, per ogni $n \in \mathbf{N}$, si ha

$$
\begin{aligned}
&\left|\frac{1}{n} \sum_{j=1}^{n-1} \mu\left(T^{-j} A \cap B\right)-\mu(A) \mu(B)\right| \\
&\leq\left|\frac{1}{n} \sum_{j=1}^{n-1} \mu\left(T^{-j} A \cap B\right)-\mu\left(T^{-j} A_0 \cap B_0\right)\right| \\
&\qquad+\left|\frac{1}{n} \sum_{j=1}^{n-1} \mu\left(T^{-j} A_0 \cap B_0\right)-\mu\left(A_0\right) \mu\left(B_0\right)\right| \\
&\qquad+\left|\mu\left(A_0\right) \mu\left(B_0\right)-\mu\left(A_0\right) \mu(B)\right|+\left|\mu\left(A_0\right) \mu(B)-\mu(A) \mu(B)\right| \\
&\leq 2 \varepsilon+\left|\frac{1}{n} \sum_{j=1}^{n-1} \mu\left(T^{-j} A_0 \cap B_0\right)-\mu\left(A_0\right) \mu\left(B_0\right)\right|+2 \varepsilon,
\end{aligned}
$$

che dimostra la (a).

(b) Analogamente a quanto fatto sopra, si ottiene, per ogni $j \in \mathbf{Z}_{+}$

$$
\begin{aligned}
&\left|\mu\left(T^{-j} A \cap B\right)-\mu(A) \mu(B)\right| \\
&\leq\left|\mu\left(T^{-j} A \cap B\right)-\mu\left(T^{-j} A_0 \cap B_0\right)\right| \\
&\qquad+\left|\mu\left(T^{-j} A_0 \cap B_0\right)-\mu\left(A_0\right) \mu\left(B_0\right)\right| \\
&\qquad+\left|\mu\left(A_0\right) \mu\left(B_0\right)-\mu\left(A_0\right) \mu(B)\right|+\left|\mu\left(A_0\right) \mu(B)-\mu(A) \mu(B)\right| \\
&\leq 4 \varepsilon+\left|\mu\left(T^{-j} A_0 \cap B_0\right)-\mu\left(A_0\right) \mu\left(B_0\right)\right| ;
\end{aligned}
$$

di qui, prendendo la media aritmetica, scende l'asserto. L'ultima sequenza di diseguaglianze stabilisce anche la (c). □

Possiamo sfruttare il corollario appena dato per stabilire quando la traslazione di Markov bilatera sia ergodica.

TEOREMA 9.3. *Per una traslazione di* MARKOV *bilatera-*(p, Π) T *sono equivalenti le affermazioni:*

(a) T *è ergodica;*
(b) *la matrice di transizione* Π *è ergodica (da ogni stato si può raggiungere un qualsiasi altro stato).*

DIMOSTRAZIONE. (a) $\Longrightarrow$ (b) Si indichi con e_j la funzione indicatrice del cilindro

$$\{\{x_n : n \in \mathbf{Z}\} : x_0 = j\}.$$

Allora, per il corollario 8.1, si ha q.c.

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{s=0}^{n-1} e_j\,(T^s\,\omega) = \int e_j\, d\mu = p_j > 0.$$

Moltiplicando ambo i membri per e_i, integrando ed usando il teorema di convergenza dominata si ottiene

$$\frac{1}{n} \sum_{s=0}^{n-1} p_i\, p_{ij}^{(s)} = \frac{1}{n} \int e_i(\omega)\, e_j\,(T^s\,\omega)\; d\mu(\omega) \xrightarrow[n \to +\infty]{} p_i\, p_j.$$

Pertanto la successione $\{\Pi^n\}$ converge nel senso di CESÀRO ad una matrice limite $A = (a_{ij})$, gli elementi della quale sono dati, quali che sia lo stato $i \in \{0, 1, \dots, k-1\}$, da

$$a_{ij} = p_j > 0.$$

Quali che siano gli stati i e j si ha

$$\frac{1}{n} \sum_{s=0}^{n-1} p_{ij}^{(s)} \xrightarrow[n \to +\infty]{} a_{ij} = p_j > 0,$$

sicché esiste un numero naturale s tale che $p_{ij}^s > 0$. Dunque la matrice Π è ergodica.

(b) $\Longrightarrow$ (a) Se Π è ergodica può essere regolare o ciclica ([**36**]). In entrambi casi la successione $\{\Pi^n : n \in \mathbf{N}\}$ converge nel senso di CESÀRO ad una matrice limite A che ha tutte le righe eguali e tutti gli elementi strettamente positivi. Per dimostrare che T è ergodica basta, in virtú del Teorema precedente, far vedere che, per ogni scelta di due cilindri A e B con

$$A = \{\{x_n : n \in \mathbf{Z}\} : x_a = i_0, x_{a+1} = i_1, \dots, x_{a+r} = i_r\}$$
$$B = \{\{x_n : n \in \mathbf{Z}\} : x_b = j_0, x_{b+1} = j_1, \dots, x_{b+r} = j_s\}$$

è

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{h=0}^{n-1} \mu\left(T^{-h}\, A \cap B\right) = \mu(A)\, \mu(B).$$

Ora, se $h > b + s - a$ si ha

$$\mu\left(T^{-h} A \cap B\right) = p_{j_0} p_{j_0 j_1} \dots p_{j_{s-1} j_s} p_{j_s i_0}^{a+h-b-s} p_{i_0 i_1} \dots p_{i_{r-1} i_r}$$

e quindi

$$\begin{aligned} \lim_{n\to+\infty} \frac{1}{n} \sum_{h=0}^{n-1} \mu\left(T^{-h} A \cap B\right) &= p_{j_0} p_{j_0 j_1} \dots p_{j_{s-1} j_s} p_{i_0} p_{i_0 i_1} \dots p_{i_{r-1} i_r} \\ &= \mu(A)\, \mu(B), \end{aligned}$$

sicché T è ergodica. □

Il seguente risultato appartiene all'analisi reale "elementare", ma consente di porre in forma differente la proprietà di mescolanza debole.

TEOREMA 9.4. *Per una successione limitata* $\{x_n : n \in \mathbf{N}\}$ *di numeri reali sono equivalenti le condizioni:*

(a) *la successione* $\{|x_n| : n \in \mathbf{Z}_+\}$ *converge a* 0 *nel senso di* CESÀRO,

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} |x_j| = 0;$$

(b) *esiste un sottoinsieme* J *di* $\mathbf{Z}_+$ *di densità nulla, tale, cioè, che*

$$\lim_{n\to+\infty} \frac{1}{n} \operatorname{card}\left[J \cap \{0, 1, \dots, n-1\}\right] = 0,$$

per il quale risulta

$$\lim_{\substack{n\to+\infty \\ n \notin J}} x_n = 0;$$

(c) *vale*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} |x_j|^2 = 0.$$

DIMOSTRAZIONE. Per ogni sottoinsieme F di $\mathbf{Z}_+$, si ponga, per semplicità,

$$\alpha_F(n) := \operatorname{card}\left[F \cap \{0, 1, \dots, n-1\}\right].$$

(a) $\Longrightarrow$ (b) Per ogni $k \in \mathbf{N}$ sia

$$J_k := \left\{ n \in \mathbf{Z}_+ : |x_n| \geq \frac{1}{k} \right\}.$$

Pertanto

$$J_1 \subset J_2 \subset \dots \subset J_n \subset \dots . \tag{9.4}$$

Inoltre, ogni J_k ha densità nulla, perché

$$\frac{1}{n}\sum_{j=0}^{n-1}|x_j| \geq \frac{1}{n}\,\frac{1}{k}\,\alpha_{J_k}(n).$$

Esistono perciò numeri naturali $0 \leq s_0 < s_1 < \dots < s_n < \dots$ tali che per $n \geq s_k$ si abbia

$$\frac{1}{n}\,\alpha_{J_{k+1}}(n) < \frac{1}{k+1}.$$

Si ponga

$$J := \cup_{k\in \mathbf{Z}_+}\{J_{k+1}\cap [s_k, s_{k+1}[\}\,;$$

Ora dalla (9.4) a da $s_k \leq n < s_{k+1}$ segue che

$$\begin{aligned} &J\cap\{0,1\dots,n-1\}\\ &\quad= [J_{k-1}\cap\{0,1,\dots,s_k-1\}]\cup[J_k\cap\{s_k,s_k+1,\dots,n-1\}]\\ &\quad\subset [J_{k-1}\cap\{0,1,\dots,s_k-1\}]\cup[J_k\cap\{0,1,\dots,n-1\}]\,, \end{aligned}$$

e, perciò,

$$\begin{aligned} \frac{1}{n}\,\alpha_J(n) &\leq \frac{1}{n}\,\left\{\alpha_{J_k}(s_k)+\alpha_{J_{k+1}}(n)\right\} \leq \frac{1}{n}\,\left\{\alpha_{J_k}(n)+\alpha_{J_{k+1}}(n)\right\}\\ &< \frac{1}{k-1}+\frac{1}{k}. \end{aligned}$$

Quando n tende a $+\infty$, anche k tende a $+\infty$; di qui segue che J ha densità nulla. Se $n \geq s_k$ e $n \notin J$ scende che n non appartiene a J_{k+1} e perciò,

$$|x_n| < \frac{1}{k+1}$$

che dà l'asserto (b).

(b) $\Longrightarrow$ (a) Si supponga che sia $|x_n| \leq H$ per ogni $n \in \mathbf{Z}_+$. Fissato $\varepsilon > 0$, esistono due numeri naturali $n_1 = n_1(\varepsilon) \in \mathbf{N}$ e $n_2 = n_2(\varepsilon) \in \mathbf{N}$ tali che sia $|x_n| < \varepsilon$ per ogni naturale $n \geq n_1$ con $n \notin J$ e

$$\frac{1}{n}\,\alpha_J(n) < \varepsilon,$$

per ogni $n \geq n_2$. Posto $n_0 := \max\{n_1, n_2\}$, per ogni $n \geq n_0$, è

$$\begin{aligned} \frac{1}{n}\sum_{j=0}^{n-1}|x_j| &= \frac{1}{n}\,\left[\sum_{j\in J\cap\{0,1,\dots,n-1\}}|x_j| + \sum_{j\notin J\cap\{0,1,\dots,n-1\}}|x_j|\right]\\ &< \frac{H}{n}\,\alpha_J(n)+\varepsilon < (H+1)\,\varepsilon, \end{aligned}$$

onde l'asserto.

(a) $\Longleftrightarrow$ (c) Basta osservare che, alla luce di quanto è stato dimostrato sinora si ha equivalenza tra le due condizioni

$$\lim_{\substack{n\to+\infty\\ n\notin J}} |x_n| = 0 \qquad \text{e} \qquad \lim_{\substack{n\to+\infty\\ n\notin J}} |x_n|^2 = 0.$$

□

Sono ora evidenti i seguenti corollari.

COROLLARIO 9.1. *Per una trasformazione $T : \Omega \to \Omega$ che conserva la misura su uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ sono equivalenti le affermazioni:*

(a) *T è debolmente mescolante;*
(b) *per ogni coppia A e B di insiemi misurabili esiste $J(A, B)$ sottoinsieme di $\mathbf{N}$ di densità nulla tale che*

$$\lim_{\substack{n\to+\infty\\ n\notin J(A,B)}} \mu\left(T^{-n} A \cap B\right) = \mu(A)\,\mu(B).$$

(c) *per ogni coppia A e B di insiemi misurabili esiste $J(A, B)$ sottoinsieme di $\mathbf{N}$ di densità nulla tale che*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \mu\left(T^{-j} A \cap B\right) - \mu(A)\,\mu(B)\right|^2 = 0.$$

Avremo modo di usare il seguente lemma.

LEMMA 9.1. *L'unione di due insiemi di densità nulla è ancora di densità nulla.*

DIMOSTRAZIONE. Siano J_1 e J_2 due sottoinsiemi di $\mathbf{N}$ di densità nulla,

$$\lim_{n\to+\infty} \frac{1}{n} \operatorname{card}\left[J_j \cap \{0, 1, \ldots, n-1\}\right] = 0 \qquad (j = 1, 2).$$

Allora

$$\begin{aligned} 0 &\le \frac{1}{n} \operatorname{card}\left[(J_1 \cup J_2) \cap \{0, 1, \ldots, n-1\}\right] \\ &= \frac{1}{n} \operatorname{card}\left[(J_1 \cap \{0, 1, \ldots, n-1\}) \cup (J_2 \cap \{0, 1, \ldots, n-1\})\right] \\ &\le \frac{1}{n} \operatorname{card}\left(J_1 \cap \{0, 1, \ldots, n-1\}\right) \\ &\qquad + \frac{1}{n} \operatorname{card}\left(J_2 \cap \{0, 1, \ldots, n-1\}\right) \xrightarrow[n\to+\infty]{} 0. \end{aligned}$$

□

10. Espressione funzionale di ergodicità e mescolanza

In questa sezione ricorreremo sistematicamente all'uso dell'operatore U_T indotto dalla trasformazione T, introdotto nella sezione 6.

TEOREMA 10.1. *Per una trasformazione $T : \Omega \to \Omega$ che conserva la misura su uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ sono equivalenti le affermazioni:*

(a) *T è ergodica;*
(b) *per coppia f e g di funzioni di L^2 è*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \langle U_T^j f, g\rangle = \langle f, 1\rangle \langle 1, g\rangle; \tag{10.1}$$

(c) *per ogni funzione f di L^2 è*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \langle U_T^j f, f\rangle = \langle f, 1\rangle \langle 1, f\rangle; \tag{10.2}$$

TEOREMA 10.2. *Per una trasformazione $T : \Omega \to \Omega$ che conserva la misura su uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ sono equivalenti le affermazioni:*

(a) *T è debolmente mescolante;*
(b) *per coppia f e g di funzioni di L^2 è*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Big| \langle U_T^j f, g\rangle - \langle f, 1\rangle \langle 1, g\rangle \Big| = 0; \tag{10.3}$$

(c) *per ogni funzione f di L^2 è*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Big| \langle U_T^j f, f\rangle - \langle f, 1\rangle \langle 1, f\rangle \Big| = 0; \tag{10.4}$$

(d) *per ogni funzione f di L^2 è*

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Big| \langle U_T^j f, f\rangle - \langle f, 1\rangle \langle 1, f\rangle \Big|^2 = 0; \tag{10.5}$$

TEOREMA 10.3. *Per una trasformazione $T : \Omega \to \Omega$ che conserva la misura su uno spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ sono equivalenti le affermazioni:*

(a) *T è fortemente mescolante;*
(b) *per coppia f e g di funzioni di L^2 è*

$$\lim_{n\to+\infty} \langle U_T^n f, g\rangle = \langle f, 1\rangle \langle 1, g\rangle; \tag{10.6}$$

(c) *per ogni funzione f di L^2 è*

$$\lim_{n\to+\infty} \langle U_T^n f, f\rangle = \langle f, 1\rangle \langle 1, f\rangle; \tag{10.7}$$

I tre teoremi appena enunciati si dimostrano in maniera molto simile. Pertanto basterà dare la dimostrazione di uno solo di essi, per esempio dell'ultimo.

Dimostrazione del Teorema 10.3. (b) $\Longrightarrow$ (a) Nella (10.6) basta prendere $f = 1_A$ e $g = 1_B$ con $A, B \in \mathcal{F}$.

(a) $\Longrightarrow$ (c) Per ogni coppia A e B di insiemi misurabili si ha

$$\begin{aligned}\lim_{n\to+\infty} \langle U_T^n 1_A, 1_B\rangle &= \lim_{n\to+\infty} \int 1_A \circ T^n\, 1_B\, d\mu = \lim_{n\to+\infty} \int_{T^{-n} A\cap B} d\mu \\ &= \lim_{n\to+\infty} \mu\left(T^{-n} A\cap B\right) = \mu(A)\,\mu(B) = \langle 1_A, 1\rangle \langle 1, 1_B\rangle.\end{aligned}$$

Se in quest'ultima relazione si tiene fisso B, si ha, per la linearità del prodotto interno rispetto al primo fattore,

$$\lim_{n\to+\infty} \langle U_T^n h, 1_B\rangle = \langle h, 1\rangle \langle 1, 1_B\rangle,$$

per ogni funzione semplice h. Si tenga ora fissa la funzione h, sicché dall'ultima relazione scende

$$\lim_{n\to+\infty} \langle U_T^n h, h\rangle = \langle h, 1\rangle \langle 1, h\rangle;$$

pertanto la (10.7) vale per la funzioni semplici.

Sia, ora f una funzione di L^2 e si fissi, arbitrariamente, $\varepsilon > 0$. Si può scegliere una funzione semplice h in modo che sia $\|f - h\|_2 < \varepsilon$. Esiste anche un numero naturale $n_0 = n_0(\varepsilon)$ tale che, per ogni $n \ge n_0$, sia

$$|\,\langle U_T^n h, h\rangle - \langle h, 1\rangle \langle 1, h\rangle\,| < \varepsilon.$$

Per ogni $n \ge n_0$, si ha

$$\begin{aligned}&|\,\langle U_T^n f, f\rangle - \langle f, 1\rangle \langle 1, f\rangle\,| \\ &\le |\,\langle U_T^n f, f\rangle - \langle U_T^n h, f\rangle\,| + |\,\langle U_T^n h, f\rangle - \langle U_T^n h, h\rangle\,| \\ &\qquad + |\langle U_T^n h, h\rangle - \langle h, 1\rangle \langle 1, h\rangle\,| + |\langle h, 1\rangle \langle 1, h\rangle - \langle f, 1\rangle \langle 1, h\rangle\,| \\ &\qquad + |\langle f, 1\rangle \langle 1, h\rangle - \langle f, 1\rangle \langle 1, f\rangle\,| \\ &= |\,\langle U_T^n (f-h), f\rangle\,| + |\,\langle U_T^n h, f-h\rangle\,| + \varepsilon + |\,\langle 1, h\rangle\,|\,|\,\langle h-f, 1\rangle\,| \\ &\qquad + |\,\langle f, 1\rangle\,|\,|\,\langle 1, h-f\rangle\,| \\ &\le \|f-h\|_2\, \|f\|_2 + \|h\|_2\, \|h-f\|_2 + \varepsilon + \|h\|_2\, \|h-f\|_2 + \|f\|_2\, \|h-f\|_2 \\ &\le \varepsilon\, \{1 + 2\, \|h\|_2 + 2\, \|f\|_2\},\end{aligned}$$

onde l'asserto.

(c) $\Longrightarrow$ (b) Sia f una funzione di L^2 e sia $\mathcal{H}_f$ il piú piccolo sottospazio chiuso di L^2 che contiene f e le funzioni costanti e che sia invariante rispetto a U_T, $U_T\,\mathcal{H}_f \subset \mathcal{H}_f$. Posto

$$\mathcal{G}_f := \left\{\varphi \in L^2 : \lim_{n\to+\infty} \langle U_T^n\, f, \varphi\rangle = \langle f, 1\rangle\, \langle 1, \varphi\rangle\right\},$$

si ha

$$\mathcal{H}_f \subset \mathcal{G}_f \subset L^2.$$

Infatti, per definizione, $\mathcal{G}_f$ è un sottospazio di L^2 e contiene f per ipotesi. Inoltre contiene la funzioni costanti

$$\begin{aligned}\langle U_T^n\, f, c\rangle &= \overline{c} \int U_T^n\, f\, d\mu = \overline{c} \int f \circ T^n\, d\mu = \overline{c} \int f\, d\left(\mu \circ T^{-n}\right)\\ &= \overline{c} \int f\, d\mu = \langle f, 1\rangle\, \langle 1, c\rangle.\end{aligned}$$

Infine $\mathcal{G}_f$ è invariante rispetto a U_T; infatti se $\varphi \in \mathcal{G}_f$ si ha

$$\begin{aligned}\lim_{n\to+\infty} \langle U_T^n\, f, U_T\, \varphi\rangle &= \lim_{n\to+\infty} \int (f \circ T^n)\, \overline{(\varphi\circ T)}\, d\mu\\ &= \lim_{n\to+\infty} \int \left(f \circ T^{n-1}\right) \overline{\varphi}\, d\mu = \lim_{n\to+\infty} \langle U_T^{n-1}\, f, \varphi\rangle\\ &= \langle f, 1\rangle\, \langle 1, \varphi\rangle = \langle f, 1\rangle \int \overline{\varphi}\, d\mu\\ &= \langle f, 1\rangle \int \overline{\varphi\circ T}\, d\mu = \langle f, 1\rangle\, \langle 1, U_T\, \varphi\rangle,\end{aligned}$$

sicché anche $U_T\, \varphi$ appartiene a $\mathcal{G}_f$.

Se φ è una funzione ortogonale al sottospazio $\mathcal{H}_f$, $\varphi \perp \mathcal{H}_f$, si ha

$$\langle U_T^n\, f, \varphi\rangle = 0$$

per ogni $n \in \mathbf{N}$. Inoltre $\langle 1, \varphi\rangle = 0$. Ma allora si ha $\mathcal{H}_f^{\perp} \subset \mathcal{G}_f$ che insieme all'inclusione $\mathcal{H}_f \subset \mathcal{G}_f$ dà $\mathcal{G}_f = L^2$. □

Sia $T : \Omega \to \Omega$ una trasformazione che conserva la misura sullo spazio di probabilità $(\Omega, \mathcal{F}, \mu)$. Si può definire la trasformazione $T\times T : \Omega\times\Omega \to \Omega\times\Omega$ mediante

$$(T \times T)(\omega, \omega') := (T(\omega), T(\omega'))\,.$$

La trasformazione cosí definita conserva la misura sullo spazio di probabilità $(\Omega\times\Omega, \mathcal{F}\otimes\mathcal{F}, \mu\otimes\mu)$. Infatti ciò è vero per i rettangoli misurabili (gli insiemi di $\mathcal{F} \times \mathcal{F}$) e, quindi, per le unioni disgiunte e finite di tali rettangoli. In virtú del Teorema 2.2, ciò basta a stabilire che $T \times T$ conserva la misura.

TEOREMA 10.4. *Per una trasformazione* $T : \Omega \longrightarrow \Omega$ *che conserva la misura su uno spazio di probabilità* $(\Omega, \mathcal{F}, \mu)$ *sono equivalenti le affermazioni*

(a) T *è debolemente mescolante;*
(b) $T \times T$ *è ergodica;*
(c) $T \times T$ *è debolmente mescolante.*

DIMOSTRAZIONE. (a) $\Longrightarrow$ (c) Siano A, B, C e D insiemi di $\mathcal{F}$. Per il Corollario 9.1, esistono due sottoinsiemi J_1 e J_2 di $\mathbf{N}$, entrambi di densità nulla, tali che

$$\lim_{\substack{n\to+\infty \\ n\notin J_1}} \mu\left(T^{-n} A \cap B\right) = \mu(A)\,\mu(B)$$
$$\lim_{\substack{n\to+\infty \\ n\notin J_2}} \mu\left(T^{-n} C \cap D\right) = \mu(C)\,\mu(D).$$

Di conseguenza, ricorrendo al Lemma 9.1, si ha

$$\begin{aligned}
&\lim_{\substack{n\to+\infty \\ n\notin J_1\cup J_2}} (\mu\otimes\mu)\left[(T\times T)^{-n}(A\times C)\cap(B\times D)\right] \\
&\qquad = \lim_{\substack{n\to+\infty \\ n\notin J_1\cup J_2}} \mu\left(T^{-n} A \cap B\right)\,\mu\left(T^{-n} C \cap D\right) \\
&\qquad = \mu(A)\,\mu(B)\,\mu(C)\,\mu(D) = \left[(\mu\otimes\mu)(A\times C)\right]\left[(\mu\otimes\mu)(C\times D)\right].
\end{aligned}$$

L'asserto vale dunque per i rettangoli misurabili di $\mathcal{F}\otimes\mathcal{F}$ e quindi anche per le unioni disgiunte e finite di tali rettangoli. Queste unioni formano un'algebra $\mathcal{A}$ che genera la tribú prodotto $\mathcal{F}\otimes\mathcal{F}$. Alla luce del Teorema 9.2 ciò implica l'asserto.

L'implicazione (c) $\Longrightarrow$ (b) è ovvia.

(b) $\Longrightarrow$ (a) Se A e B sono insiemi misurabili si ha

$$\begin{aligned}
\frac{1}{n}\sum_{j=0}^{n-1}\mu\left(T^{-j} A\cap B\right) &= \frac{1}{n}\sum_{j=0}^{n-1}(\mu\otimes\mu)\left[(T\times T)^{-j}(A\times\Omega)\cap(B\times\Omega)\right] \\
&\xrightarrow[n\to+\infty]{} \left[(\mu\otimes\mu)(A\times\Omega)\right]\left[(\mu\otimes\mu)(B\times\Omega)\right] \\
&= \mu(A)\,\mu(B).
\end{aligned}$$

Ma si ha anche, dall'ipotesi,

$$\begin{aligned}
\frac{1}{n}\sum_{j=0}^{n-1}\mu^2\left(T^{-j} A\cap B\right) &= \frac{1}{n}\sum_{j=0}^{n-1}(\mu\otimes\mu)\left((T\times T)^{-j}(A\times A)\cap(B\times B)\right) \\
&\xrightarrow[n\to+\infty]{} (\mu\otimes\mu)(A\times A)\,(\mu\otimes\mu)(B\times B) \\
&= \mu^2(A)\,\mu^2(B).
\end{aligned}$$

Perciò

$$\begin{aligned}
&\frac{1}{n}\sum_{j=0}^{n-1}\left\{\mu\left(T^{-j} A\cap B\right)-\mu(A)\,\mu(B)\right\}^2\\
&=\frac{1}{n}\sum_{j=0}^{n-1}\left\{\mu^2\left(T^{-j}A\cap B\right)-2\,\mu\left(T^{-j}A\cap B\right)\,\mu(A)\,\mu(B)+\mu^2(A)\,\mu^2(B)\right\}\\
&\qquad\qquad\xrightarrow[n\to+\infty]{} 2\,\mu^2(A)\,\mu^2(B)-2\,\mu^2(A)\,\mu^2(B)=0
\end{aligned}$$

che conclude la dimostrazione. □

DEFINIZIONE 10.1. Sia $T:\Omega\to\Omega$ una trasformazione che conserva la misura sullo spazio di probabilità $(\Omega,\mathcal{F},\mu)$. Si dice che il numero $k\in\mathbf{C}$ è un *autovalore* di T, se esiste una funzione $f\in L^2$ con $f\neq 0$ tale che $U_T\,f=k\,f$, o, equivalentemente, tale che $f\circ T=k\,f$ q.c.. Si dice allora che f è un'*autofunzione* corrispondente all'autovalore k.

Si noti che tutti gli autovalori di T hanno modulo unitario poiché, se vale $U_T\,f=k\,f$, è

$$\|f\|_2^2=\|U_T\,f\|_2^2=\|k\,f\|_2^2=|k|^2\,\|f\|_2^2.$$

Inoltre $k=1$ è sempre un autovalore (che corrisponde alla funzioni costanti).

DEFINIZIONE 10.2. Si dice che la trasformazione $T:\Omega\to\Omega$ che conserva la misura sullo spazio di probabilità $(\Omega,\mathcal{F},\mu)$ ha *spettro continuo* se 1 è l'unico autovalore e le funzioni q.c. costanti sono le sole autofunzioni dell'operatore U_T indotto da T.

Al prossimo risultato che lega i concetti di mescolanza debole e di spettro continuo, è opportuno premettere il seguente teorema, per la dimostrazione del quale rimando a [**27**].

TEOREMA 10.5. *Sia $U:\mathcal{H}\to\mathcal{H}$ un operatore unitario sopra uno spazio di* HILBERT *(complesso) $\mathcal{H}$. Per ogni $f\in\mathcal{H}$ esiste un'unica misura μ_f definita sui boreliani della circonferenza unitaria $K:=\{z\in\mathbf{C}:|z|=1\}$ tale che, per ogni $n\in\mathbf{N}$, sia*

$$\langle U^n\,f,f\rangle=\int_K z^n\,d\mu_f(z).$$

Si noti che se $T:\Omega\to\Omega$ è una trasformazione invertibile che conserva la misura sullo spazio di probabilità $(\Omega,\mathcal{F},\mu)$, allora l'operatore U_T indotto da T è unitario; se, inoltre, T ha spettro continuo, allora, per ogni funzione $f\in L^2$ con $\langle f,1\rangle=0$, μ_f non ha atomi.

TEOREMA 10.6. *Per una trasformazione $T:\Omega\to\Omega$ che conserva la misura sullo spazio di probabilità $(\Omega,\mathcal{F},\mu)$ sono equivalenti le affermazioni*

(a) *T è debolmente mescolante;*
(b) *T ha spettro continuo.*

DIMOSTRAZIONE. (a) $\Longrightarrow$ (b). Si supponga che T sia debolmente mescolante e, se possibile, che k sia un autovalore differente da 1, $U_t\, f = k\, f$ con $k \neq 1$. Integrando si ottiene

$$\begin{aligned}\langle f, 1\rangle &= \int f\, d\mu = \int f\, d\left(\mu \circ T^{-1}\right) = \int (f \circ T)\, d\mu = \int U_T\, f\, d\mu \\ &= \langle U_T\, f, 1\rangle = k\, \langle f, 1\rangle.\end{aligned}$$

Di qui scende $\langle f, 1\rangle = 0$, sicché, ricordando che $|k| = 1$, il Teorema 10.2 dà

$$\begin{aligned}0 &= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \langle U_T^j\, f, f\rangle \right| = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \langle k^j\, f, f\rangle \right| \\ &= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} |k|^j\, \left| \langle f, f\rangle \right| = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \langle f, f\rangle \right| = \left| \langle f, f\rangle \right|.\end{aligned}$$

Dunque $f = 0$ q.c., sicché k non può essere un autovalore. Se $k = 1$ si ha necessariamente che f è q.c. costante, perché T è ergodica.

(b) $\Longrightarrow$ (a) Supposto che T abbia spettro continuo, si vuole dimostrare che T è debolmente mescolante; basta, a tal fine, dimostrare che vale la

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \langle U_T^j\, f, f\rangle - \langle f, 1\rangle\, \langle 1, f\rangle \right|^2 = 0;$$

Se f è q.c. costante, questa è un'immediata conseguenza della diseguaglianza di SCHWARZ, che in questo caso, si riduce ad un'eguaglianza. Si può, quindi, supporre che sia $\langle f, 1\rangle = 0$. Per il Teorema 10.5, basta far vedere che, per ogni misura μ sui boreliani di K, si ha

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left| \int_K z^j\, d\mu \right|^2 = 0. \tag{10.8}$$

Ricorrendo al teorema di FUBINI si ottiene

$$\begin{aligned}&\frac{1}{n} \sum_{j=0}^{n-1} \left| \int_K z^j\, d\mu \right|^2 = \frac{1}{n} \sum_{j=0}^{n-1} \left\{ \int_K z^j\, d\mu(z) \right\} \left\{ \int_K \overline{w}^j\, d\mu(w) \right\} \\ &= \frac{1}{n} \sum_{j=0}^{n-1} \iint_{K\times K} (z\,\overline{w})^j\, d(\mu\otimes\mu)(z, w) = \iint_{K\times K} \left\{ \frac{1}{n} \sum_{j=0}^{n-1} (z\,\overline{w})^j \right\} d(\mu\otimes\mu)(z, w).\end{aligned}$$

Se il punto $(z\,\overline{w})$ non appartiene alla diagonale di $K \times K$, ovvero se $z\,\overline{w} \neq 1$, allora si ha

$$\frac{1}{n}\sum_{j=0}^{n-1}(z\,\overline{w})^j = \frac{1}{n}\,\frac{1-(z\,\overline{w})^n}{1-z\,\overline{w}} \xrightarrow[n\to+\infty]{} 0.$$

Poiché la misura μ è continua, è $(\mu\otimes\mu)(z,z)=0$, sicché l'integrando nell'ultimo integrale scritto tende a 0 q.o.; inoltre, esso è maggiorato in modulo da 1 sicché la (10.8) è ora una conseguenza del teorema di convergenza dominata. □

11. Topologie per le trasformazioni

Vogliamo qui introdurre alcune topologie sulla famiglia delle trasformazioni che conservano la misura. Ci limiteremo, seguendo HALMOS, a considerare solo un caso particolare; supporremo cosí che siano verificate le seguenti ipotesi

— si considereranno solo trasformazioni invertibili sicché l'insieme nel quale si introdurrà la topologia è, di fatto, un gruppo;
— si considereranno come coincidenti le trasformazioni che differiscono su un insieme di misura nulla; in altre parole, anziché il gruppo delle trasformazioni invertibili che conservano la misura, si considererà il gruppo degli automorfismi dell'algebra di misura associata allo spazio;
— si supporrà che lo spazio di probabilità in esame sia costiuito dall'intervallo unitario $[0,1]$ munito della tribú di BOREL $\mathcal{B}$ e della misura di LEBESGUE λ su $\mathcal{B}$.

Ricordiamo qui la definizione delle *topologie forte e debole* in uno spazio di HILBERT H. Data una famiglia di operatori su H, $\{T_n : n \in I\}$, ove I è un insieme diretto, si dice che tale famiglia *converge fortemente* all'operatore T se, per ogni $f \in H$, si ha

$$\|T_n f - Tf\| \longrightarrow 0.$$

Si dice invece che la famiglia in questione *converge debolmente* a T se, per ogni coppia f e g di elementi di H, si ha

$$\langle T_n f, g\rangle \longrightarrow \langle Tf, g\rangle.$$

È ben noto che la convergenza forte implica quella debole, senza che sia vero il viceversa. Si è visto che esiste una corrispondenza biunivoca tra gli automorfismi di un'algebra di misura e gli operatori unitarî su L^2, sicché il gruppo degli automorfismi eredita ogni topologia della famiglia degli operatori unitarî su L^2. In quest'ultimo insieme le due topologie, forte e debole, coincidono.

TEOREMA 11.1. *Per una famiglia $\{U_n : n \in I\}$ di operatori unitarî su L^2 su uno spazio di* HILBERT *H le seguenti proprietà sono equivalenti:*

(a) *$\{U_n : n \in I\}$ converge fortemente a un operatore unitario U;*
(b) *$\{U_n : n \in I\}$ converge debolmente ad un operatore unitario U.*

DIMOSTRAZIONE. Poiché la convergenza forte implica quella debole, basta dimostrare l'implicazione (b) $\Longrightarrow$ (a). Per ogni $f \in H$, si ha

$$\begin{aligned}\|U_n f - Uf\| &= \langle U_n f - Uf, U_n f - Uf\rangle \\ &= \langle U_n f, U_n f\rangle - \langle U_n f, Uf\rangle - \langle Uf, U_n f\rangle + \langle Uf, Uf\rangle \\ &= \|f\|^2 - \langle U_n f, Uf\rangle - \langle Uf, U_n f\rangle + \|f\|^2 \\ &\xrightarrow[n\to+\infty]{} 2\,\|f\|^2 - 2\,\langle Uf, Uf\rangle = 2\,\|f\|^2 - 2\,\|f\|^2 = 0\end{aligned}$$

□

Introduciamo ora la definizione di convergenza che servirà per gli scopi della teoria ergodica.

DEFINIZIONE 11.1. Si dirà che una famiglia $\{T_n : n \in I\}$ di automorfismi *converge debolmente* a T se $T_n\,A$ converge a $T\,A$ per ogni elemento A dell'algebra di misura $\widetilde{\mathcal{B}}$, vale a dire se, per ogni insieme $A \in \mathcal{B}$, è

$$\lambda\,(T_n\,A \Delta T\,A) \longrightarrow 0.$$

Una sottobase per gli aperti di questa topologia è data dagli insiemi della forma

$$N(T, A, \varepsilon) := \{S : \lambda\,(T\,A \Delta S\,A) < \varepsilon\}\,.$$

Consideriamo la famiglia degli intervalli

$$\left\{\left]\frac{k}{2^n}, \frac{k+1}{2^n}\right] : n \in \mathbf{Z}_+, k = 0, 1, \ldots, 2^n - 1\right\}.$$

Per n fissato in $\mathbf{Z}_+$, ognuno di tali intervalli si dirà *di rango n*, mentre si dirà *insieme di rango n* ogni unione, necessariamente finita, di intervalli di rango n. Si dirà *permutazione (diadica di rango n)* una trasformazione invertibile che conserva la misura e che porta ogni intervallo diadico di rango n in un altro intervallo diadico di rango n. La permutazione ciclica di rango n permuta ciclicamente gli intervalli diadici di rango n.

Per dimostrare il Teorema di approssimazione debole occorrerà un risultato preliminare.

La famiglia $\mathcal{B}_0$ di tutti gli insiemi diadici è densa nell'algeba di misura, è stabile rispetto alle operazioni booleane finite; inoltre la misura di un insieme diadico è un numero razionale diadico. Se E è un insieme di $\mathcal{B}_0$ e se $r \in [0, \lambda(E)]$ è un umero razionale, allora E ha sottoinsiemi che appartengono a $\mathcal{B}_0$ e che hanno misura eguale a r. Se E è in $\mathcal{B}_0$, anche TE gode della proprietà appena illustrata.

LEMMA 11.1. *Sia $\{E_1, E_2, \dots, E_k\}$ una partizione dell'intervallo unitario; se $\delta > 0$ è arbitrario, e se r_1, $r_2, \dots$, r_k sono k numeri razionali tali che*

$$|\lambda(E_i) - r_i| < \delta \qquad (i = 1, 2, \dots, k),$$

allora esiste una partizione $\{F_1, F_2, \dots, F_k\}$ dell'intervallo unitario in insiemi di $\mathcal{B}_0$, $F_i \in \mathcal{B}_0$ per $i = 1, 2, \dots, k$, tali che, per ogni $i = 1, 2, \dots, k$, si abbia

$$\lambda\,(E_1 \Delta F_i) < 2\,\varepsilon, \qquad e \qquad \lambda\,(F_i) = r_i.$$

DIMOSTRAZIONE. Si scelga, per ogni $i = 1, 2, \dots, k$, un insieme $B_i \in \mathcal{B}_0$ tale che

$$\lambda\,(E_i \Delta B_i) < \gamma,$$

o, equivalentemente

$$|\lambda\,(E_i) - \lambda\,(B_i)| = |\lambda\,(E_i \setminus B_i) - \lambda\,(B_i \setminus E_i)| \le \lambda\,(E_i \Delta F_i) < \gamma.$$

Di conseguenza si ha, per ogni $i = 1, 2, \dots, k$,

$$|\lambda\,(B_i) - r_i| < \gamma + \delta.$$

Gli insiemi B_i cosí costruiti non costituiscono necessariamente una partizione dell'intervallo unitario e non hanno necessariamente la misura richiesta. Poiché si ha

$$\lambda\,[(E_i \cap E_j)\,\Delta\,(B_i \cap B_j)] \le \lambda\,(E_i \Delta B_i) + \lambda\,(E_j \Delta B_j) < 2\gamma$$

e

$$\lambda\,(E_i \cap E_j) = 0,$$

segue che

$$\lambda\,[B_i \cap (\cup_{j \ne i} B_j)] < 2k\gamma.$$

Si prendano ora gli insiemi di $\mathcal{B}_0$ definiti, per $i = 1, 2, \dots, k$, da

$$B'_i := B_i \setminus \left(\bigcup_{j \ne i} B_j \right).$$

Per questi insiemi è, per $i = 1, 2, \dots, k$,

$$|\lambda\,(B'_i) - r_i| < \delta + (2k+1)\gamma \qquad \text{e} \qquad |\lambda\,(E_i) - \lambda\,(B'_i)| < (2k+1)\gamma.$$

Gli insiemi B'_i sono ora a due a due disgiunti, ma la loro unione non è ancora tutto $[0,1]$ e la misura di ciascuno di essi non è ancora quella desiderata.

Se $\lambda(B'_i) > r_i$, si consideri un sottoinsieme B''_i di B'_i appartenente a $\mathcal{B}_0$ e tale che

$$\lambda(B'_i) - \lambda(B''_i) = r_i;$$

si unisca quindi B''_i al complemento degli insiemi B'_i,

$$C := [0,1] \setminus \left(\bigcup_{i=1}^{k} B'_i \right) \in \mathcal{B}_0$$

e si ponga $F_i = B'_i \setminus B''_i$.

Se invece si ha $\lambda(B'_i) < r_i$ si prenda un sottoinsieme $B''_i \in \mathcal{B}_0$ di C tale che

$$\lambda(B'_1 \cup B''_i) = r_i$$

e si ponga $F_i := B'_i \cup B''_i$. Gli insiemi F_i $(i = 1, 2, \dots, k)$ cosí costruiti costituiscono una partizione di $[0,1]$ in insiemi di $\mathcal{B}_0$ e $\lambda(F_i) = r_i$. Per questi insiemi si ha , per $i = 1, 2, \dots, k$,

$$\lambda(E_i \Delta F_i) < \delta + 2(2k+1)\gamma.$$

Basta ora prendere γ sufficientemente piccolo perché sia

$$\lambda(E_i \Delta F_i) < 2\delta;$$

ciò conclude la dimostrazione. □

Un *intorno diadico* di un automorfismo S di G ha la forma

$$\{T \in G : \lambda(SD \Delta TD) < \varepsilon\},$$

ove D è un insieme diadico. Gli intorni diadici costiuiscono una sottobase per le topologia debole di G. Il risultato fondamentale asserisce che le permutazioni sono dense in G.

TEOREMA 11.2 (di approssimazione debole). *Ogni intorno diadico contiene permutazioni cicliche di rango arbitrario.*

DIMOSTRAZIONE. Dimostriamo dapprima che, se P è una permutazione e se

$$N := \{T : \lambda(PD_i \Delta TD_i) < \varepsilon \; i = 1, 2, \dots, n\}$$

è un intorno diadico di P, allora N contiene permutazioni cicliche di rango arbitrariamente grande. Ogni insieme PD_i è un insieme diadico. Ne segue che esiste un naturale j tale che PD_i sia di rango j. Naturalmente, j può essere scelto arbitrariamente grande; si scleg ora $k \in \mathbf{N}$ tale che sia

$$k > j \qquad \text{e} \qquad \frac{2^{j+1}}{2^k} < \varepsilon.$$

Ogni intervallo diadico di rango j è l'unione di 2^{k-j} intervalli diadici di rango k. Definiamo ora una permutazione ciclica Q di rango k che appartenga all'intorno N.

Sia I un intervallo di rango j. Se P non lascia I invariato, si definica Q in modo che mandi il primo, nell'ordine naturale, sottointervallo di rango k contenuto in I su tutto il primo sottointervallo di rango k di PI. Se P non manda PI su I, si definisca Q in modo che mandi il primo sottointervallo di rango k di PI nel primo sottointervallo di rango k di P^2I. Si continui in questo modo sino a che non si raggiunga l'ultimo elemento nel ciclo indotto da P su I, sia esso $P^{q-1}I$. Quindi $P\left(P^{q-1}I\right) = I$. Allora si definisca Q in modo che mandi il primo sottointervallo di rango k di $P^{q-1}I$ nel *secondo* sottointervallo di rango k di I e si proceda come sopra sostituendo ovunque "primo sottointervallo" con "secondo sottointervallo". Alla fine, si definisca Q in modo che mandi il *secondo* sottointervallo di rango k di $P^{q-1}I$ nel *terzo* sottointervallo di rango k di I. Si ripeta di nuovo la procedura sino a quando Q non debba essere definito sull'ultimo sottointervallo di rango k di $P^{q-1}I$.

Sia ora F un intervallo diadico di rango j disgiunto da I, PI,..., $P^{q-1}I$ e si definisca Q in modo che applichi l'ultimo sottointervallo di rango k di $P^{q-1}I$ sul primo sottointervallo di rango k di F. Si ripeta, quindi, su F l'intera procedura fatta su I e si proceda in maniera analoga sino a quando non si sia completata la definizione di Q sull'intera decomposizione di P. La definizione di Q sarà completata quando si applicherà l'ultimo sottointervallo di rango k dell'ultimo termine della decomposizione di P nel primo sottointervallo di rango k dell'intervallo originale I.

Evidentemente, Q è una permutazione ciclica di rango k. Per costruzione, se E è una arbitrario intervallo diadico di rango j, allora si ha

$$\lambda\,(PE\Delta QE) < \frac{2}{2^k}.$$

Ma un insieme diadico è l'unione disgiunta di al piú 2^j intervalli diadici di rango j, sicché se D è un insieme diadico di rango j, allora è

$$\lambda\,(PD\Delta QD) < 2^j\,\frac{2}{2^k} = \frac{2^{j+1}}{2^k} < \varepsilon.$$

Perciò Q appartiene a N.

Rimane da dimostrare che le permutazioni sono dense in G. A tal fine, sia T un'arbitraria trasformazione e sia

$$N := \{S : \lambda\,(SD_i\Delta TD_i) < \varepsilon\ i = 1, 2, \dots, n\}$$

un intorno diadico di T; vogliamo dimostrare che N contiene una permutazione. Gli insiemi $D_i \cap TD_j$ costituiscono una partizione dell'intervallo unitario. Sia $\delta > 0$ tale che

$$n^2\,\delta < \frac{\varepsilon}{2}.$$

Il Lemma 11.1 assicura l'esistenza di una partizione $\{E_{ij}\}$ dell'intervallo unitario in sottoinsiemi diadici tale che sia

$$\lambda\left((D_i \cap TD_j)\,\Delta E_{ij}\right) < \delta.$$

Una seconda applicazione dello stesso lemma assicura l'esistenza di una partizione $\{F_{ij}\}$ dell'intervallo unitario in sottoinsiemi diadici tale che sia

$$\lambda\left((D_i \cap TD_j)\,\Delta TF_{ij}\right) < \delta$$

e che risulti, per tutti gli indici $\lambda(E_{ij}) = \lambda(F_{ij})$. poiché T conserva la misura, si ha anche

$$\lambda\left((D_i \cap TD_j)\,\Delta F_{ij}\right) < \delta.$$

Sia k un numero naturale tale che tutti gli insiemi D_i, E_{ij} e F_{ij} abbiano rango k e sia P una permutazione di rango k che applichi F_{ij} su E_{ij}. Poiché

$$D_j = \cup_i \left(T^{-1}D_i \cap D_j\right),$$

si ha

$$\lambda\left(D_j\Delta \cup_i F_{ij}\right) < n^2\,\delta < \frac{\varepsilon}{2}.$$

Ma allora

$$\lambda\left(PD_j\Delta \cup_i E_{ij}\right) < \frac{\varepsilon}{2}.$$

Pertanto

$$\lambda\left(PD_j\Delta TD_j\right) \leq \lambda\left(PD_j\Delta \cup_i E_{ij}\right) + \lambda\left(\cup_i E_{ij}\Delta TD_j\right) < \varepsilon,$$

ciò che conclude la dimostrazione. □

Vi è una seconda topologia sul gruppo G degli automorfismi, detta *topologia uniforme.*

Si introducono su G due metriche

$$d_1(S,T) := \sup\left\{\lambda\left(SA\Delta TA\right) : A \in \mathcal{B}(G)\right\}$$
$$d_2(S,T) := \lambda\{S\omega \neq T\omega\}.$$

È immediato controllare, ed è lasciato come esercizio, che tanto d_1 quanto d_2 sono metriche su G. Inoltre, esse sono invarianti rispetto alle operazioni di gruppo (moltiplicazione a destra o a sinista, e formazione dell'inversa); ciò è ovvio per d_1. Per quanto riguarda d_2, è immediata l'invarianza per la moltiplicazione a sinistra. Poiché si ha

$$S\left\{S\omega \neq T\omega\right\} = \left\{S^{-1}\omega \neq T^{-1}\omega\right\}$$

segue che

$$\begin{aligned} d_2\left(S^{-1}, T^{-1}\right) &= \lambda\left(\{S^{-1}\omega \neq T^{-1}\omega\}\right) \\ &= \lambda\left(S\{S\omega \neq T\omega\}\right) = \lambda\left(\{S\omega \neq T\omega\}\right) = d_2(S,T), \end{aligned}$$

ciò che mostra l'invarianza di d_2 rispetto alla formazione dell'inversa. L'invarianza di d_2 rispetto alla moltiplicazione a destra è ora conseguenza dell'invarianza rispetto alla moltiplicazione a sinistra e di quella rispetto alla formazione dell'inversa.

Se per $x \in [0,1]$ e per $n \in \mathbf{N}$, si ha $T^n x = x$, si dice che T è *periodica in* x; il minimo $n = n(x) \in \mathbf{N}$ tale che $T^n x = x$ si dice il *periodo di* x. Se $T^n = I$ (I è l'identità su $[0,1]$), allora si dice che T è periodica e il minomo $n \in \mathbf{N}$ tale che $T^n = I$ si chiama *periodo di* T. Evidentemente se Tè periodica è periodica per ogni x; il viceversa non è necessariamente vero. Se la trasformanzione T nonè periodica in alcum punto di $[0,1]$, si dirà che è *aperiodica*. Sia A_n l'insieme di punti di $[0,1]$ nei quali T è periodica di periodo $n \in \mathbf{N}$,

$$A_n := \left\{x \in [0,1] : T^n x = x, T^j x \neq x (j < n)\right\}, \tag{11.1}$$

e sia A_0 l'insieme dei punti nei quali T è antiperiodica,

$$A_0 := \left(\cup_{n\in\mathbf{N}} A_n\right)^c = \cap_{n\in\mathbf{N}} A_n^c. \tag{11.2}$$

LEMMA 11.2. *Se T è periodica di periodo n per quasi ogni punto di $[0,1]$, allora esiste un boreliano E di $[0,1]$ di misura $1/n$ tale che gli insiemi E, TE,..., $T^{n-1}E$ siano a due a due disgiunti.*

DIMOSTRAZIONE. Se $n = 1$ non vi è nulla da dimostrare. Se $n > 1$ esiste necessariamente un boreliano E_1 tale che

$$\lambda\left(E_1 \Delta T E_1\right) > 0.$$

Infatti, se per ogni boreliano E_1 di $[0,1]$ fosse $\lambda\left(E_1 \Delta T E_1\right) = 0$, T sarebbe di periodo 1 per quasi ogni punto. Poiché T conserva la misura si ha

$$\begin{aligned} \lambda\left(E_1 \setminus T E_1\right) &= \lambda\left(E_1\right) - \lambda\left(E_1 \cap T E_1\right) \\ &= \lambda\left(T E_1\right) - \lambda\left(E_1 \cap T E_1\right) = \lambda\left(T E_1 \setminus E_1\right), \end{aligned}$$

segue che $\lambda\left(E_1 \setminus T E_1\right) > 0$. Posto

$$F_1 := E_1 \setminus T E_1,$$

l'insieme F_1 è un boreliano di misura positiva che, per costruzione è disgiunto da $T F_1 = T E_1 \setminus T E_1$. Se $n = 2$, la dimostrazione è terminata. Se $n > 2$, esiste necessariamente un sottoinsieme E_2 di F_1 che è misurabile e tale che $\lambda\left(E_2 \Delta T^2 E_2\right) > 0$; se cosí non fosse, T sarebbe di periodo 2 quasi ovunque in F_1. Posto $F_2 := E_2 \setminus T^2 E_2$, scende come sopra che che F_2 e $T^2 F_2$ sono disgiunti. Procedendo in questo modo per induzione, dopo $n-1$ passi si sono costruiti gli insiemi F_1, F_2, ..., F_{n-1} di misura positiva tali che $F_1 \supset F_2 \supset \cdots \supset F_{n-1}$ e che, per ogni $j = 1, 2, \ldots, n-1$ gli insiemi F_j e $T^j F_j$ siano disgiunti. Posto $E := F_{n-1}$, l'insieme E risulta essere disgiunto da $T^j E$ per ogni $j = 1, 2, \ldots, n-1$. La misura di E potebbe non essere eguale a $1/n$; si appliche allora lo stesso procedimanto ora usato al sottoinsieme invariante

$$\left(E \cup TE \cup \cdots \cup T^{n-1}E\right)^c.$$

Si proceda per induzione terminando al piú dopo infinità numerabile di passi. $\square$

LEMMA 11.3. *Se T è antiperiodica, allora per ogni $n \in \mathbf{N}$ e per ogni $\varepsilon > 0$, esiste un boreliano E tale che gli insiemi E, TE,..., $T^{n-1}E$ siano disgiunti e che sia*

$$\lambda\left(E \cup TE \cup \cdots \cup T^{n-1}E\right) > 1 - \varepsilon.$$

DIMOSTRAZIONE. Sia p un numero naturale tale che $1/p < \varepsilon$. Ricorrendo allo stesso ragionamento del Lemma 11.2 nel quale si sostituisce pn a n, si costrusice un boreliano F di misura positiva e tale che gli insiemi F, TF,..., $T^{pn-1}F$ siano a due a due disgiunti. Inoltre F è massimale nel senso che non esiste alcun boreliano che includa F, abbia misura maggiore di quest'ultimo ed abbia le proprietà elencate per F. Perciò, se F_0 è un sottoinsieme misurabile di $T^{pn-1}F$ con $\lambda(F_0) > 0$, allora esiste almeno un numero naturale $j = 1, 2, \ldots, pn$ tale che $\lambda\left(T^j F_0 \cap F\right) > 0$, altrimenti l'insieme $F \cup TF_0$ contraddirrebbe la massimalità di F. Si definisca ora, per $j = 1, 2, \ldots, pn$, il boreliano

$$F_j := \left(T^{pn-1}F \cap T^{-j}F\right) \bigcap \left(\cup_{i=1}^{j-1} T^{-i}F^c\right).$$

Allora i ha

$$\lambda\left(T^{pn-1}F \setminus \cup_{j=1}^{pn} F_j\right) = 0.$$

Gli insiemi di una stessa colonna dello schema

$$\begin{array}{llll}
TF_2 & & & \\
TF_3 & T^2F_3 & & \\
\vdots & \vdots & & \\
TF_{pn} & T^2TF_{pn} & \ldots & T^{pn-1}TF_{pn}
\end{array} \tag{11.3}$$

sono a due a due digiunti perché si ottengono trasformando mediante la stessa potenza di T insiemi disgiunti; inoltre, due insiemi di colonne differenti sono disgiunti perché le loro immagini inverse mediante opportune potenze di T sono incluse in insiemi distinti di F, TF, ..., $T^{pn-1}F$.

Gli insiemi $T^i F_j$ $(1 \leq i < j \leq pn)$ sono disgiunti da ognuno degli insiemi $T^k F$ $(k = 1, 2, \ldots, pn-1)$ e quindi dalla loro unione. Infatti, se è $k > i$, allora

$$T^i F_j \cap T^k F \subset T^i \left(T^{pn-1}F \cap T^{k-i}F\right) = \emptyset;$$

se, invece, $k \leq i$, allora si avrebbe $0 \leq i - k < j$, sicché, per la definizione di F_j, è vuoto l'insieme

$$T^{i-k} F_j \cap F;$$

perciò è vuoto anche l'insieme

$$T^i F_j \cap T^k F = T^k \left(T^{i-k} F_j \cap F\right).$$

Pertanto gli insiemi TF_1, $T^2 F_2$, ..., $T^{pn} F_{pn}$ sono sottoinsiemi a due a due disgiunti di F; quindi

$$\lambda\left(\cup_{j=1}^{pn} T^j F_j\right) = \sum_{j=1}^{pn} \lambda\left(T^j F_j\right) = \sum_{j=1}^{pn} \lambda\left(F_j\right) = \lambda\left(T^{pn-1} F\right).$$

Si può ora asserire che l'insieme

$$F^* := \left(\cup_{k=0}^{pn-1} T^k F\right) \bigcup \left(\cup_{1 \leq i < j \leq pn} T^i F_j\right)$$

è quasi invariante, nel senso che $\lambda\left(F^* \Delta T^{-1} F^*\right) = 0$. Siccome T è antiperiodica, si ha

$$\lambda\left(F^*\right) = 1,$$

perché altrimenti l'insieme $F \cup (F^*)^c$ permetterebbe di ingrandire l'insieme massimale F. Si ponga, infine

$$E := \left(\cup_{k=0}^{p-1} T^{kn} F\right) \bigcup \left(\cup_{i=1}^{pn-1} \cup_{j=i+n}^{pn} T^i F_j\right).$$

Per costruzione gli insiemi E, TE, ..., $T^{n-1} E$ sono disgiunti. L'unione di questi insiemi non contine al piú $n-1$ insiemi di ogni riga dello schema (11.3). Ciò significa che la misura totale degli insiemi della j–esima riga non inclusi nell'unione $E \cup TE \cup \cdots \cup T^{n-1} E$ ha misura minore di $n\,\lambda\left(F_j\right)$; di conseguenza

$$\lambda\left(E \cup TE \cup \cdots \cup T^{n-1} E\right) > 1 - \sum_{j=1}^{pn} n\,\lambda\left(F_j\right) = 1 - n\,\lambda(F).$$

Gli insiemi F, TF, ..., $T^{pn-1} F$ sono disgiunti ed hanno tutti la stessa misura sicché

$$n\,\lambda(F) \leq n\,\frac{1}{pn} = \frac{1}{p} < \varepsilon.$$

Ciò conclude la dimostrazione □

Le topologie indotte su G dalle metriche d_1 e d_2 sono equivalenti; piú precisamente vale il seguente

Lemma 11.4. *Valgono le diseguaglianze*

$$\frac{2}{3}\, d_2 \leq d_1 \leq d_2.$$

DIMOSTRAZIONE. Poiché tanto d_1 quanto d_2 sono invarianti rispetto alle operazioni di gruppo, basta dimostrare che

$$\frac{2}{3}\, d_2(I,T) \leq d_1(I,T) \leq d_2(I,T)$$

per ogni trasformazione T che conservi la misura. Se

$$F := \{x \in [0,1] : Tx \neq x\},$$

allora si ha $d_2(I,T) = \lambda(F)$. Si osservi che l'insieme F è invariante rispetto a T e che tale è anche ogni sottoinsieme misurabile di F^c. Per ogni boreliano E di $[0,1]$, si ha

$$\begin{aligned} E\Delta TE &= [(E\Delta TE) \cap F] \bigcup [(E\Delta TE) \setminus F] \\ &= [(E \cap F)\,\Delta\,(TE \cap F)] \bigcup [(E \setminus F)\,\Delta\,(TE \setminus F)] \\ &= [(E \cap F)\,\Delta\,(TE \cap TF)] \bigcup [(E \setminus F)\,\Delta\,(TE \setminus TF)] \\ &= [(E \cap F)\,\Delta T\,(E \cap F)] \bigcup [(E \setminus F)\,\Delta T\,(E \setminus F)]\,; \end{aligned}$$

di qui, dato che $T\,(E \setminus F) = E \setminus F$,

$$\begin{aligned} \lambda\,(E\Delta TE) &= \lambda\,[(E \cap F)\,\Delta T\,(E \cap F)] + \lambda\,[(E \setminus F)\,\Delta T\,(E \setminus F)] \\ &= \lambda\,[(E \cap F)\,\Delta\,(TE \cap F)] = \lambda\,[(E\Delta TE) \cap F] \leq \lambda(F). \end{aligned}$$

Ne segue che $d_1(I,T) \leq d_2(I,T)$.

Si consideri ora la partizione di $[0,1]$ nei sottoinsiemi

$$\{A_0, A_1, \dots, A_n, \dots\}$$

delle equazioni (11.1) e (11.2). Si applichi il lemma 11.3 all'insieme A_0 con $n = 2$ e $\varepsilon \leq 1/3$. Esiste perciò un boreliano E_0 di A_0 tale che $E_0 \cap TE_0 = \emptyset$ e che

$$\lambda\,(E_0) \geq \frac{1}{3}\,\lambda\,(A_0)\,.$$

Applicando il Lemma 11.2 ad ogni A_n si ha un boreliano E_n di A_n tale che gli insiemi E_n, TE_n,..., $T^{n-1}E_n$ siano a due a due disgiunti e che sia

$$\lambda\,(E_n) = \frac{1}{n}\,\lambda\,(A_n)\,.$$

Si definisca ora una successione di insiemi misurabili $\{B_n : n = 0 \text{ e } n \geq 2\}$ come segue. Si ponga

— per $n = 0$, $B_0 := E_0$;
— per $n \geq 2$ pari,

$$B_n := E_n \bigcup \left(\cup_{k=1}^{n/2-1} T^{2k} E_n\right);$$

— per $n \geq 2$ dispari,

$$B_n := E_n \bigcup \left(\cup_{k=1}^{(n-3)/2} T^{2k} E_n \right).$$

Si ha, in ogni caso, $B_n \cap TB_n = \emptyset$. Se $n \geq 2$ è pari, si ha

$$\lambda(B_n) = \frac{1}{2} \lambda(A_n),$$

mentre, se $n \geq 2$ è dispari, si ha

$$\lambda(B_n) = \frac{1}{2} \left(1 - \frac{1}{n}\right) \lambda(A_n).$$

Ne segue che, per $n \geq 2$, è

$$\lambda(B_n) \geq \frac{1}{3} \lambda(A_n).$$

Posto

$$B := F_0 \cup B_2 \cup B_3 \cup \cdots \cup B_n \cup \ldots,$$

si ha $B \cap TB = \emptyset$ e

$$\lambda(B) \geq \frac{1}{3} (1 - \lambda(A_1)).$$

Quindi

$$d_1(I, T) \geq \lambda(B \Delta TB) \geq \frac{2}{3} (1 - \lambda(A_1)) = \frac{2}{3} d_2(I, T),$$

ciò che conclude la dimostrazione. □

Per ottenere il teorema di approssimazione per la topologia uniforme avremo bisogno del seguente risultato che interessse indipendente.

TEOREMA 11.3. *Per ogni coppia A e B di boreliani dell'intervallo* $[0, 1]$ *esiste una trasformazione* $T : [0, 1] \to [0, 1]$ *tale che* $\lambda(TA \Delta B) = 0$.

DIMOSTRAZIONE. Se i due boreliani hanno misura nulla, $\lambda(A) = \lambda(B) = 0$, non vi è nulla da dimostrare. Supporremo, pertanto, che sia $\lambda(A)\,\lambda(B) > 0$. Sia S una trasformazione ergodica su $[0, 1]$; esiste allora un naturale n tale che

$$\lambda(S^n A \cap B) > 0.$$

Si definisca $Tx := S^n x$ se x appartiene a $A \cap T^{-n} B$. Si proceda ora per induzione.

Se $\lambda(A \setminus S^{-n} B) > 0$, si applichi il ragionamento appena usato per trovare un sottoinsieme misurabile C di $A \setminus S^{-n} B$ di misura non nulla, $\lambda(C) > 0$ ed un naturale k tale che

$$\lambda\left[S^k C \cap (B \setminus S^{-n} A)\right] > 0;$$

e si definisca $Tx := S^k x$ se x appartiene a C. Procedendo in questo modo, eventualmente per induzione transfinita, si giunge a definire la trasformazione T su tutto A. Applicando poi lo stesso metodo ai due insiemi A^c e B^c, si definisce T su tutto $[0,1]$. □

Il seguente teorema asserisce di fatto che la trasformazioni periodiche sono dense in G nella topologia uniforme.

TEOREMA 11.4 (d'approssimazione uniforme). *Per ogni trasformazione aperiodica T, per ogni numero naturale n e per ogni $\varepsilon > 0$ esiste una trasformazione S di periodo n tale che*

$$d_2(S,T) \leq \frac{1}{n} + \varepsilon.$$

DIMOSTRAZIONE. In virtú del Lemma 11.3 esiste un boreliano E tale che gli insiemi E, TE,..., $T^{n-1}E$ siano a due a due disgiunti e che

$$\lambda\left(E \cup TE \cup \cdots \cup T^{n-1}E\right) > 1 - \varepsilon.$$

Se x è in $E \cup TE \cup \cdots \cup T^{n-2}E$, si definisca $Sx := Tx$; se x appartiene a $T^{n-1}E$, si ponga $Sx := T^{-n+1}x$. La trasformazione S è cosí stata definita in $E \cup TE \cup \cdots \cup T^{n-1}E$ ed è qui di periodo n; quale che sia il modo di estendere la defınzione di S a tutto $[0,1]$, si ha

$$d_2(S,T) \leq \lambda\left(T^{n-1}E\right) + \varepsilon = \lambda(E) + \varepsilon \leq \frac{1}{n} + \varepsilon.$$

Rimane quindi solo da dimostrare che si può definire S nel complementare di $E \cup TE \cup \cdots \cup T^{n-1}E$ in modo che anche qui S sia di periodo n e questo segue facilmente dal Teorema 11.3. □

12. Categoria

L'introduzione della topologia debole nell'insieme delle trasformazioni che conservano la misura consente di trarre conclusioni importanti sulla grandezza topologica di alcuni sottinsiemi di trasformazioni.

TEOREMA 12.1. *Nella topologia debole le trasformazioni fortemente mescolanti costituiscono un insieme di prima categoria.*

DIMOSTRAZIONE. Per $k \in \mathbf{N}$ sia $\mathcal{P}_k$ l'insieme delle trasformazioni che conservano la misura e che sono periodiche di periodo k,

$$\mathcal{P}_k := \left\{T \in G : T^k = I\right\};$$

si ponga inoltre

$$\mathcal{P}_n^* := \cup_{k>n} \mathcal{P}_k.$$

In virtú del Teorema di approssimazione debole, l'insieme $\mathcal{P}_n^*$ è denso per ogni $n \in \mathbf{N}$.

Posto $E := [0, 1/2]$ si consideri la seguente famiglia di trasformazioni che conservano la misura

$$M_k := \left\{ T \in G : \left| \lambda \left(E \cap T^k E \right) - \frac{1}{4} \right| \leq \frac{1}{5} \right\}.$$

Tale insieme è chiuso nella topologia debole. Si definiscano

$$M_n^* := \cap_{k>n} M_k \quad \text{e} \quad M^* := \cup_{n\in\mathbf{N}} \cap_{k>n} M_k = \liminf_{n\to+\infty} M_k.$$

È evidente che M^* contiene tutte le trasformazioni fortemente mescolanti; infatti, per queste si ha

$$\lim_{n\to+\infty} \lambda \left(E \cap T^k E \right) = \frac{1}{4}.$$

Basta, pertanto, dimostrare che M^* è di prima categoria; a tal fine basta dimostrare che, per ogni $n \in \mathbf{N}$, M_n^* è e, siccome M_n^* è chiuso, basta in effetti mostrare che il suo complementare in G, $G \setminus M_n^*$ è denso in G. Ma

$$G \setminus M_n^* = G \cap \left(\cap_{k>n} M_k \right)^c = \cup_{k>n} G \setminus M_k,$$

sicché basta dismostrare che vale l'inclusione

$$\mathcal{P}_k \subset G \setminus M_k.$$

Questo è però di facile dimostrazione, perché se T appartiene a $\mathcal{P}_k$ si ha $T^k E = E$, sicché

$$\lambda \left(E \cap T^k E \right) - \frac{1}{4} = \lambda(E) - \frac{1}{4} = \frac{1}{4};$$

dunque T non appartiene a M_k. □

Il lemma che segue sarà necessario per studiare l'insieme delle trasformazioni debolmente mescolanti.

Lemma 12.1. *Nella topologia debole la classe coniugata di ogni trasformazione antiperiodica T_0,*

$$\left\{ S^{-1} T_0 S : S \in G \right\},$$

è densa in G.

Dimostrazione. Sia $N = \{T \in G : \lambda \left(P D_i \Delta T D_i \right) < \varepsilon, \ i = 1, 2, \ldots, n\}$ un intorno diadico di una permutazione P; si vuole dimostrare che esiste una trasformazione S di G tale $S^{-1} T_0 S$ appartenga a N. Posto

$$M = \left\{ T \in G : \lambda \left(P D_i \Delta T D_i \right) < \frac{\varepsilon}{2}, \ i = 1, 2, \ldots, n \right\},$$

segue dal Teorema di approssimazione debole che l'intorno M contiene una permutazione ciclica Q di rango k maggiore dei ranghi di tutti gli insiemi D_i $(i = 1, 2, \ldots, n)$ e tale che

$$\frac{1}{2^{k-2}} < \varepsilon.$$

Il Teorema di approssimazione uniforme assicura ora l'esistenza di una trasformazione $R \in G$ che ha ovunque periodo 2^k e tale che

$$d_2(R, T_0) < \frac{2}{2^k} < \frac{\varepsilon}{2}. \tag{12.1}$$

Le trasformazioni Q e R sono coniugate in G. Si numerino gli intervalli diadici di rango k, E_0, $E_1, \ldots$, E_q $(q = 2^k)$ in modo tale che sia $QE_i = E_{i+1}$, con gli indici presi modulo q. Sia ora F_0 un boreliano tale che $\lambda(F_0) = 1/q$ e che gli insiemi F_0, $RF_0, \ldots$, $R^{q-1}F_0$ siano a due a due disgiunti. Posto $F_i := R^iF_0$ $(i = 1, 2, \ldots, q-1$, sia S una qualsiasi trasformazione che conserva la misura e che manda E_0 su F_0. Si può estendere la definizione di tale S a tutto l'intervallo unitario in maniera ricorsiva, ponendo, per $x \in E_i$ $(i = 1, 2, \ldots, q-1)$,

$$Sx := R^iSQ^{-i}x.$$

Schematicamente si ha

$$\begin{array}{ccccccccccc}
E_0 & \xrightarrow{Q} & E_1 & \xrightarrow{Q} & E_2 & \xrightarrow{Q} & \ldots & \xrightarrow{Q} & E_{q-2} & \xrightarrow{Q} & E_{q-1} \\
{\scriptstyle S}\big\downarrow & & {\scriptstyle S}\big\downarrow & & {\scriptstyle S}\big\downarrow & & {\scriptstyle S}\big\downarrow & & {\scriptstyle S}\big\downarrow & & {\scriptstyle S}\big\downarrow \\
F_0 & \xrightarrow[R]{} & F_1 & \xrightarrow[R]{} & F_2 & \xrightarrow[R]{} & \ldots & \xrightarrow[R]{} & F_{q-2} & \xrightarrow[R]{} & F_{q-1}
\end{array}$$

Si vede subito che $Q = S^{-1}RS$.

Poiché la distanza d_2 è invariante per le operazioni di gruppo in G, la (12.1) dà

$$d_2\left(Q, S^{-1}T_0S\right) = d_2\left(S^{-1}RS, S^{-1}T_0S\right) = d_2(R, T_0) < \frac{\varepsilon}{2}.$$

Pertanto

$$\begin{aligned}
\lambda\left(PD_i\Delta S^{-1}T_0SD_i\right) &\leq \lambda(PD_i\Delta QD_i) + \lambda\left(QD_i\Delta S^{-1}T_0SD_i\right) \\
&\leq \lambda(PD_i\Delta QD_i) + d\left(Q, S^{-1}T_0S\right) \\
&< \frac{\varepsilon}{2} + d_2\left(Q, S^{-1}T_0S\right) < \varepsilon.
\end{aligned}$$

Perciò $Q, S^{-1}T_0S$ appartiene a N. □

13. Trasformazioni indotte

DEFINIZIONE 13.1. Siano $T\Omega \to \Omega$ una traformazione che conserva la misura sullo spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ e A un sottoinsieme misurabile di Ω.

Sia $\omega \in A$ un punto ricorrente (quasi tutti i punti di A per il teorema di ricorrenza di POINCARÉ sono ricorrenti, ritornano, cioè, in A) e si indichi con $n_A(\omega)$ il minimo naturale positivo tale che $T^{n_A(\omega)}\,\omega$ sia in A,

$$n_A(\omega) := \{n \in \mathbf{N} : T^n\omega \in A\}.$$

Si dirà *trasformazione indotta da T su A* la trasformazione $T_A : A \to A$ definita per quasi tutti i punti di A da

$$T_A\,\omega := T^{n_A(\omega)}\,\omega.$$

Si noti che, a rigore, T_A non è definita su tutti i punti di A, ma solo su quelli ricorrenti; sui rimanenti punti T_A può essere definita in maniera arbitraria.

Si considerino la traccia della tribú $\mathcal{F}$ su A,

$$\mathcal{F}_A := \mathcal{F} \cap A = \{A \cap B : B \in \mathcal{F}\}$$

e la probabilità μ_A condizionata da A,

$$\mu_A(B) = \frac{\mu\,(A \cap B)}{\mu(A)}.$$

TEOREMA 13.1. *La trasformazione indotta T_A è misurabile rispetto alla tribú traccia $\mathcal{F}_A$ e conserva la misura sullo spazio di probabilità $(A, \mathcal{F}_A, \mu_A)$.*

DIMOSTRAZIONE. Per ogni $n \in \mathbf{N}$ si definiscano gli insiemi

$$\begin{aligned} A_n &:= \left\{\omega \in A : T\omega \notin A, T^2\omega \notin A, \dots, T^{n-1}\omega \notin A, T^n\omega \in A\right\} \\ &= \{\omega \in A : n_A(\omega) = n\} \end{aligned}$$

e

$$B_n := \left\{\omega \in A^c : T\omega \notin A, T^2\omega \notin A, \dots, T^{n-1}\omega \notin A, T^n\omega \in A\right\}.$$

Ora si ha

$$A_n = \left(A \cap T^{-n}\,A\right) \setminus \cup_{j=1}^{n-1}\left(A \cap T^{-j}\,A\right) \quad \text{e} \quad B_n = T_n\,A \setminus \cup_{j=0}^{n-1} T^{-j}\,A,$$

sicché la misurabilità di T assicura che tanto A_n quanto B_n siano misurabili, per ogni $n \in \mathbf{N}$. Inoltre, per definizione, si ha $A_n \subset A$ ($n \in \mathbf{N}$); infine, se B è un sottoinsieme misurabile di A, $B \in \mathcal{F}_A$, è

$$T_A^{-1}\,B = \cup_{n\in\mathbf{N}}\left(A_n \cap T^{-n}\,B\right), \tag{13.1}$$

come subito si verifica. Questo mostra che T_A è misurabile rispetto a $\mathcal{F}_A$.

Gli insiemi della successione $\{A_n :\in \mathbf{N}\}$ sono disgiunti; inoltre

$$\mu\,(A \setminus \cup_{n\in\mathbf{N}} A_n) = 0.$$

D'altro canto si ha

$$T^{-1}\,A = A_1 \cup B_1$$

e, per ogni $n \in \mathbf{N}$,

$$T^{-1} B_n = A_{n+1} \cup B_{n+1}.$$

Pertanto, se B appartiene a $\mathcal{F}_A$

$$T^{-1} B = T^{-1} (B \cap A) = (A_1 \cap T^{-1} B) \cup (B_1 \cap T^{-1} B) \tag{13.2}$$

$$T^{-1} (B_n \cap T^{-n} B) = \left(A_{n+1} \cap T^{-(n+1)} B\right) \cup \left(B_{n+1} \cap T^{-(n+1)} B\right). \tag{13.3}$$

Poiché T conserva la misura su $(\Omega, \mathcal{F}, \mu)$, per ogni $n \in \mathbf{N}$, si ha dalla ultime relazioni applicate ripetutemante

$$\begin{aligned}\mu(B) &= \mu\left(T^{-1} B\right) = \mu\left(A_1 \cap T^{-1} B\right) + \mu\left(B_1 \cap T^{-1} B\right) \\ &= \mu\left(A_1 \cap T^{-1} B\right) + \mu\left(A_2 \cap T^{-2} B\right) + \mu\left(B_2 \cap T^{-2} B\right) = \dots \\ &= \sum_{j=1}^{n} \mu\left(A_j \cap T^{-j} B\right) + \mu\left(B_n \cap T^{-n} B\right).\end{aligned}$$

La successione $\{B_n : n \in \mathbf{N}\}$ è formata da insiemi disgiunti e, di conseguenza, ha come limite $\emptyset$; perciò

$$\mu(B) = \sum_{n \in \mathbf{N}} \mu\left(A_n \cap T^{-n} B\right).$$

Ma per la (13.1) quest'ultima somma è eguale a $\mu\left(T_A^{-1} B\right)$, onde $\mu(B) = \mu\left(T_A^{-1} B\right)$. □

Teorema 13.2. *Se T è invertibile, anche la trasformazione indotta T_A è invertibile.*

Dimostrazione. Poiché T^{-1} conserva la misura, si può definire la trasformazione indotta $\left(T^{-1}\right)_A$ su A da T^{-1}. Vogliamo mostrare che $T_A^{-1} = \left(T^{-1}\right)_A$. A tal fine basta far vedere che si ha $\left(T^{-1}\right)_A T_A \omega = \omega$ per quasi tutti i punti ω di A. Se ω è in A_n allora $T_A \omega = T^n \omega =: y \in A$. Chiaramente $T^{-n} y = \omega \in A$. Si supponga che sia $z = T^{-m} y \in A$ per $m \in \mathbf{N}$ con $1 \leq m < n$. Allora $T^m z = y = T^m (T^{n-m} \omega)$. Ma $T^{n-m} \omega$ non può appartenere a A, ciò che contraddice all'ipotesi che T sia iniettiva. Ne segue che $\left(T^{-1}\right)_A y = T^{-n} y = \omega$. □

Teorema 13.3. *Se T è ergodica, tale è anche T_A.*

Dimostrazione. Sia B un insieme invariante di $\mathcal{F}_A$, $T^{-1} B = B$. Si ponga

$$D := \cup_{n \in \mathbf{N}} \left\{\left(A_n \cap T^{-n} B\right) \cup \left(B_n \cap T^{-n} B\right)\right\}.$$

poiché $T^{-1} B = B$, la (13.1) dà $B \supset A_n \cap T^{-n}\mathcal{B}$ $(n \in \mathbf{N})$ e, quindi, per ogni $n \in \mathbf{N}$, si ha

$$A \setminus B = A \cap B^c \subset A \cap \left(A_n^c \cup T^{-n} B^c\right) = (A \cap A_n^c) \cup \left(A \cap T^{-n} B^c\right)$$

e, dunque,

$$\begin{aligned} A \setminus B &\subset \cap_{n\in\mathbf{N}} (A \cap A_n^c) \cup \left(A \cap T^{-n} B^c\right) \\ &= \{\cap_{n\in\mathbf{N}} (A \cap A_n^c)\} \cup \left\{\cap_{n\in\mathbf{N}} \left(A \cap T^{-n} B^c\right)\right\} \\ &= (A \cap A^c) \cup \left\{\cap_{n\in\mathbf{N}} \left(A \cap T^{-n} B^c\right)\right\} = \cap_{n\in\mathbf{N}} \left(A \cap T^{-n} B^c\right). \end{aligned}$$

D'altro canto

$$D^c = \cap_{n\in\mathbf{N}} \left\{(A_n \cup B_n) \cap T^{-n} B\right\}^c = \cap_{n\in\mathbf{N}} \left\{(A_n \cup B_n)^c \cup T^{-n} B^c\right\},$$

sicché $A \setminus B$ è contenuto in D^c. Ora, la (13.2) e la (13.3) danno

$$T^{-1} \left(A_n \cap T^{-n} B\right) \subset T^{-1} B = \left(A_1 \cap T^{-1} B\right) \cup \left(B_1 \cap T^{-1} B\right)$$

e

$$T^{-1} \left(B_n \cap T^{-n} B\right) = \left(A_{n+1} \cap T^{-(n+1} B\right) \cup \left(B_{n+1} \cap T^{-(n+1)} B\right),$$

sicché $T^{-1} D = D$, cioè D è invariante. Poiché T è ergodica, si ha o $\mu(D) = 0$ o $\mu(D^c) = 0$. Nel primo caso è

$$\mu(D) \geq \sum_{n\in\mathbf{N}} p\left(A_n \cap T^{-n} B\right) = \mu\left(T_A^{-1} B\right),$$

e, quindi, $\mu(B) = \mu\left(T_A^{-1} B\right) = 0$. Se, invece, $\mu(D^c) = 0$, allora

$$\mu(D^c) \geq \mu(A \setminus B) = \mu_A(B^c)\, \mu(A),$$

vale a dire $\mu_A(B^c) = 0$. □

Sia dato uno spazio mensurale $(\Omega, \mathcal{F}, \mu)$ finito o σ–finito. e una successione $\{B_n : n \in \mathbf{Z}_+\}$ di insiemi misurabili e disgiunti tali che $\mu(B_{n+1}) \leq \mu(B_n) < +\infty$. Si supponga che, per ogni $n \in \mathbf{Z}_+$, esista una trasformazione $\sigma_n : B_{n+1} \longrightarrow B_n$ invertibile e che conserva la misura tra B_{n+1} e $\sigma_n(B_{n+1})$. Sia, infine, $T_0 : B_0 \to B_0$ una trasformazione suriettiva e invertibile che conserva la misura. Posto

$$\Omega' := \cup_{n\in\mathbf{Z}_+} B_n$$

si definisca un'applicazione $T : \Omega' \longrightarrow \Omega'$ come segue

$$T\,\omega := \begin{cases} \sigma_n^{-1}\omega, & \omega \in \sigma_n(B_{n+1}), \\ T_0\,\sigma_0 \ldots \sigma_{n-1}\,\omega, & \omega \in B_n \setminus \sigma_n(B_{n+1}). \end{cases} \tag{13.4}$$

TEOREMA 13.4. *La trasformazione $T : \Omega' \longrightarrow \Omega'$ definita dalla* (13.4) *è invertibile e conserva la misura. Se, poi, è $\mu(\Omega') = 1$, allora T_0 è la trasformazione indotta da T su B_0.*

DIMOSTRAZIONE. Si supponga che sia $T\,\omega_1 = T\,\omega_2$. Varî casi sono possibili.

Caso 1: ω_1 e ω_2 appartengono a $\sigma\,(B_{n+1})$ per lo stesso indice n. Allora, $T\,\omega_1 = T\,\omega_2$ significa $\sigma_n^{-1}\,\omega_1 = \sigma_n^{-1}\,\omega_2$, e, quindi, in virtú dell'invertibilità di σ_n, $\omega_1 = \omega_2$.

Caso 2: ω_1 e ω_2 appartengono a $B_n \setminus \sigma\,(B_{n+1})$ per lo stesso indice n. In questo caso si ha

$$T\,\omega_1 = T_0\,\sigma_0 \ldots \sigma_{n-1}\,\omega_1 = T_0\,\sigma_0 \ldots \sigma_{n-1}\,\omega_2 = T\,\omega_2,$$

onde $\omega_1 = \omega_2$, perché le trasformazionei T_0, σ_0, $\ldots$, σ_{n-1} sono tutte invertibili.

Caso 3: $\omega_1 \in \sigma_n\,(B_{n+1})$ e $\omega_2 \in B_j \setminus \sigma_j\,(B_{j+1})$; ma, allora, $T\,\Omega_1$ appartiene a B_{n+1} mentre $T\,\omega_2$ è in B_0, di modo che nonsi può avere l'eguaglianza $T\,\omega_1 = T\,\omega_2$. Dunque questo caso non si può presentare.

Caso 4: è palesemente impossibile che sia $\omega_1 \in \sigma_i\,(B_{i+1})$ e $\omega_2 \in \sigma_j\,(B_{j+1})$ con $i \neq j$.

Caso 5: infine, si potrebbe, *a priori*, avere $\omega_1 \in B_i \setminus \sigma_i\,(B_{i+1})$ e $\omega_2 \in B_j \setminus \sigma_j\,(B_{j+1})$ con $i \neq j$. Si supponga, senza perdita di generalità, che sia $i < j$.. Dall'eguaglianza $T\,\omega_1 = T\,\omega_2$ scende, a causa dell'invertibilità di tutte le trasformazioni in gioco, che $\omega_1 = \sigma_i\,\sigma_{i+1} \ldots \sigma_{j-1}\,\omega_2$ che appartiene a $\sigma_i\,(B_{i+1})$ contrariamente all'ipotesi; dunque anche questo caso non si può presentare.

Si è cosí stabilito che T, definita dalla (13.4) è iniettiva. Sia ora y un punto di Ω' e si supponga che esso appartenga a B_n con $n \in \mathbf{N}$. Poiché σ_{n-1} è, per ipotesi, invertibile, esiste sicuramente un punto z di $\sigma_{n-1}\,(B_n)$ tale che sia $y = \sigma_n^{-1}\,z$, o, equivalentemente $y = T\,z$. Se, poi, y appartiene a B_0, esiste, in virtú dell'inveritibilità di T_0, $y' \in B_0$ tale che $y = T_0\,y'$. Ora, sey' è in $B_0 \setminus \sigma_0\,(B_1)$ basta prendere $z = y'$ per avere $y = T_0\,y' = T\,z$. Se, invece y' è in $\sigma_0\,(B_1)$, esiste $y_1 \in B_1$ tale che $\sigma_1\,y_1 = y'$; se y_1 è in $B_1 \setminus \sigma_1\,(B_2)$, si prenda $z = y_1$ e si avrà $T\,z = T_0\,\sigma_0\,y_1 = y$. procedendo per induzione, se y_1 appartiene a $\sigma_1\,(B_2)$, esiste $y_2 \in B_2$ tale che $\sigma_1\,y_2 = y_1$. Se per un indice $n \in \mathbf{N}$ si ha $y_n \in B_n \setminus \sigma_n\,(B_{n+1})$, si prenda $z = y_n$ e si avrà cosí $T\,z = T_0\sigma_o \ldots \sigma_{n-1}\,y_n = y$. Si supponga, quindi, che y_n appartenga a $\sigma_n\,(B_{n+1})$ per ogni $n \in \mathbf{N}$. poiché tutte le trasformazioni T_0, σ_0, $\ldots$,

$\sigma_n, \ldots$ conservano la misura e poiché $\mu(B_n) \xrightarrow[n\to+\infty]{} 0$, risulta che i punti y che danno origine a una tale successione costituiscono un insieme di misura nulla.

Per mostrare che T conserva la misura sia A un sottoinsieme misurabile di Ω' e si scriva

$$A = \bigcup_{n\in\mathbf{Z}_+} \{[A \cap \sigma_n(B_{n+1})] \cup [A \cap (B_n \setminus \sigma_n(B_{n+1}))]\}$$

che è un'unione disgiunta. Ora

$$\begin{aligned} &T^{-1} A \\ &= \bigcup_{n\in\mathbf{Z}_+} \left\{\sigma_n\, [A \cap \sigma_n(B_{n+1})] \cup \sigma_{n-1}^{-1} \ldots \sigma_0^{-1} T_0^{-1} [A \cap (B_n \setminus \sigma_n(B_{n+1}))]\right\} \end{aligned}$$

che mostra, intanto, che $T^{-1} A$ è misurabile. Inoltre

$$\begin{aligned} \mu\left(T^{-1} A\right) &= \sum_{n\in\mathbf{Z}_+} \mu\left(\sigma_n\, [A \cap \sigma_n(B_{n+1})]\right) \\ &\quad + \sum_{n\in\mathbf{Z}_+} \mu\left(\sigma_{n-1}^{-1} \sigma_{n-2}^{-1} \ldots \sigma_0^{-1} T_0^{-1}\, [A \cap (B_n \setminus \sigma_n(B_{n+1}))]\right) \\ &= \sum_{n\in\mathbf{Z}_+} \mu(A \cap \sigma_n(B_{n+1})) + \sum_{n\in\mathbf{Z}_+} \mu(A \cap (B_n \setminus \sigma_n(B_{n+1}))) \\ &= \mu(A). \end{aligned}$$

Resta da mostrare che $T_0 = T_{B_0}$. Per definizione,

$$T_{B_0}(\omega) = T^{n_{B_0}(\omega)}(\omega).$$

Se $\omega \in B_0 \setminus \sigma_0(B_))$, è $n_{B_0}(\omega) = 1$ e, dunque, $T_{B_0}(\omega) = T(\omega) = T_0(\omega)$. Se, invece, $\omega \in \sigma_0(B_1)$, si ha $T(\omega) \in B_1$; se, ora, $T(\omega)$ appartiene a $B_1 \setminus \sigma_1(B_2)$, è $n_{B_0}(\omega) = 2$ e

$$T_{B_0}(\omega) = T^2(\omega) = T(T(\omega)) = T\left(\sigma_0^{-1}(\omega)\right) = T_0\left(\sigma_0\, \sigma_0^{-1}(\omega)\right) = T_0(\omega).$$

Si supponga che $T^k(\omega)$ sia in $B_k \setminus \sigma_k(B_{k+1})$, con $T^j(\omega) \in \sigma_j(B_{j+1})$ ($j = 0, 1, \ldots, k-1$). Allora è $n_{B_0}(\omega) = k+1$ e

$$\begin{aligned} T_{B_0}(\omega) = T^{k+1}(\omega) = T\left(T^k(\omega)\right) &= T_0\, \sigma_0 \ldots \sigma_{k-1}\left(T^k(\omega)\right) \\ &= T_0\, \sigma_0 \ldots \sigma_{k-2}\, \sigma_{k-1}\, \left[T\left(T^{k-1}(\omega)\right)\right] \\ &= T_0\, \sigma_0 \ldots \sigma_{k-2}\, \left(\sigma_{k-1}\, \sigma_{k-1}^{-1}\right)\left(T^{k-1}(\omega)\right) \\ &= \cdots = T_0\left(\sigma_0\, \sigma_0^{-1}(\omega)\right) = T_0(\omega). \end{aligned}$$

I pinti $\omega \in B_0$ tali che $T^n(\omega)$ appartenga a $\sigma_n(B_{n+1})$ per ogni $n \in \mathbf{Z}_+$ costituiscono un insieme di misura nulla. $\square$

14. Esempî di trasformazioni mescolanti

Incominciamo dando l'esempio di una trasformazione ergodica che non è totalmente ergodica.

Sia Ω un insieme costituito da due punti $\Omega = \{\omega_1, \omega_2\}$ e sulla famiglia delle parti di Ω si definisca una misura di probabilità μ mediante $\mu(\{\omega_j\}) = p_j$ con $p_j = 1/2$ $(j = 1, 2)$. La trasformazione $T : \Omega \longrightarrow \Omega$ definita da $T\omega_1 = \omega_2$ e $T\omega_2 = \omega_1$, conserva la misura ed è invertibile; essa è ergodica perché i soli insiemi invarianti sono $\emptyset$ e Ω, ma non è totalmente ergodica, perché, per esempio, tanto $\{\omega_1\}$ quanto $\{\omega_2\}$ sono invarianti rispetto a T^2, che, perciò, non è invariante.

Per contro, se a non è una radice dell'unità, la rotazione T_a della circonferenza unitaria è totalmente ergodica, perché $T_a^k z = a^k z$ e neanche a^k è una radice dell'unità, sicché T_a^k è ergodica quale che sia $k \in \mathbf{N}$.

Riesiaminiamo qui gli esempî di traformazioni ergodiche della sezione 1.5.

Esempio 14.1. L'identità I sullo spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ è fortemente mescolante se, e solo e se, è ergodica, vale a dire quando tutti gli insiemi di $\mathcal{F}$ hanno probabilità eguale a 0 o a 1.

Esempio 14.2. Nessuna rotazione T_a della circonferenza unitaria $\mathcal{K}$ con $a \neq 1$ può essere debolmente mescolante. Infatti, si consideri la funzione $\varphi : \mathcal{K} \longrightarrow \mathcal{K}$ definita da $\varphi(z) := z$. Per questa si ha

$$(U_{T_a}\varphi)(z) = \varphi(T_a z) = T_a z = az = a\,\varphi(z),$$

sicché a è un autovalore e φ un'autofunzione di U_{T_a}; basta, ora, far riferimento al Teorema 10.6.

Esempio 14.3. La traslazione bilatera è fortemente mescolante Siano A e B due cilindri

$$\begin{aligned} A &= \{\, \{x_n : n \in \mathbf{Z}\} : x_a = i(0), x_{a+1} = i(1), \dots, x_{a+r} = i(r)\} \\ B &= \{\, \{x_n : n \in \mathbf{Z}\} : x_b = j(0), x_{b+1} = j(1), \dots, x_{b+r} = j(s)\}\,. \end{aligned}$$

Pur di prendere n sufficientemente grande, piú precisamente, se $n > b+s-a$, si ha

$$\mu\left(T^{-n} A \cap B\right) = p_{i(0)}\, p_{i(0)} \dots p_{i(r)}\, p_{j(0)}\, p_{j(1)} \dots p_{j(s)} = \mu(A)\,\mu(B).$$

Il Teorema 10.6 dà quindi l'asserto.

Analogamente si fa vedere che è fortemente mescolante la traslazione unilatera.

TEOREMA 14.1. *Per una traslazione bilatera di* MARKOV–(p, Π) T *sono equivalenti le affermazioni:*

(a) T *è fortemente mescolante;*
(b) *la matrice di transizione* Π *è regolare, vale a dire che esiste un numero naturale* K *tale che la matrice* Π^K *abbia tutti gli elementi positivi,* $\Pi^K > 0$.

DIMOSTRAZIONE. (a) $\Longrightarrow$ (b) Si può procedere come nella dimostrazione del Teorema 9.3 per verificare che si ha, quali che siano gli stati i e j,

$$\lim_{n \to +\infty} p_{ij}^{(n)} = p_j > 0,$$

sicché, per n sufficientemente grande, è $p_{ij}^{(n)} > 0$.

L'implicazione inversa (b) $\Longrightarrow$ (a) si dimostra come nel Teorema appena richiamato. □

Non daremo l'esempio di una trasformazione debolemente mescolante che non sia fortemente mescolante, benché, come si sia visto, tali trasformazioni debbano essere numerose. Per una tale trasformazione si veda l'esempio di KAKUTANI riportato in [**53**, pp. 87–89].

15. Misure invarianti

Si supponga che sia dato uno spazio mensurale $(\Omega, \mathcal{F}, \mu)$, ove la misura μ è σ–finita, e che $T : \Omega \longrightarrow \Omega$ sia una trasformazione misurabile, che per semplicità si supporrà anche invartibile. Il problema è di trovare una misura ν definita sullo spazio misurabile $(\Omega, \mathcal{F})$ tale che la trasformazione T sia invariante rispetto a ν. Il problema, in questa generalità non ha molto senso, perché, ad esempio, la misura identicamente nulla, soddisfà alla proprietà richiesta. Per eliminare queste misure invarianti banali, si richiede che la misura data μ sia assolutamente continua rispetto a ν, $\mu \ll \nu$. Tuttavia, anche cosí il problema ammette soluzioni non significative; basta, per esempio prendere come misura ν la misura del contare, che è ovviamente una misura invariante. A questo inconveniente si rimedia chiedendo che ν sia σ–finita.

Vi sono due problemi legati all'esistenza delle misure invarianti: nelle condizioni specificate

(a) esiste una misura *finita* ν tale che la misura data μ sia assolutamente continua rispetto a ν e che ν risulti invariante rispetto alla trasformazione T?

(b) esiste una misura σ–finita ν con le stesse proprietà?

Senza perdita di generalità si può supporre che la misura data μ sia finita. Infatti, perché μ è σ–finita esiste una partizione numerabile di Ω $\{E_n : n \in \mathbf{N}\}$ in insiemi disgiunti e misurabili. Sia $\{\alpha_n\}$ una successione di numeri reali positivi, $\alpha_n > 0$ $(n \in \mathbf{N})$, tali che sia convergente ls serie $\sum_{n\in\mathbf{N}} \alpha_n \, \mu(E_n)$. Si definisca una misura μ_0 su $(\Omega, \mathcal{F})$ mediante

$$\mu_0(E) := \sum_{n\in\mathbf{N}} \alpha_n \, \mu\left(E \cap E_n\right).$$

La misura μ_0 cosí introdotta è finita ed è equivalente a μ (cioè le due misure μ_0 e μ si annullano sugli stessi insiemi, o, equivalentemente, sono l'una assolutamente continua rispetto all'altra, $\mu \ll \mu_0$ e $\mu_0 \ll \mu$). Pertanto, μ è assolutamente continua rispetto a ν se, e solo e se, tale è anche μ_0. Sostituendo μ con μ_0 si può, quindi, supporre che μ sia una misura di probabilità $\mu(\Omega) = 1$.

Si può anche supporre, di nuovo senza perdita di generalità, che la trasformazione T non sia *singolare*; una trasformazione (invertibile) si dice singolare se esiste un insieme misurabile E tale che $\mu(E) = 0$, mentre almeno uno tra gli insiemi TE e $T^{-1}E$ ha misura diversa da zero, $\mu\left(T^{-1}E\right) \vee \mu(TE) > 0$. Sia $\{\alpha_n : n \in \mathbf{Z}\}$ una successione di numeri reali positivi, $\alpha_n > 0$ $(n \in \mathbf{Z})$, tali che $\sum_{n\in\mathbf{Z}} \alpha_n = 1$. Posto

$$\mu_1(E) := \sum_{n\in\mathbf{Z}} \alpha_n \, \mu\left(T^n E\right),$$

si controlla immediatamente che μ_1 è una misura di probabilità e che μ è assolutamente continua rispetto a μ_1, $\mu \ll \mu_1$. Se μ è assolutamente continua rispetto ad una misura invariante ν, $\mu \ll \nu$, allora anche anche μ_1 è assolutamente continua rispetto a ν, $\mu_1 \ll \nu$; infatti, se fosse $\nu(E) = 0$, allora per l'invarianza rispetto a T si avrebbe $\nu\left(T^n E\right) = 0$ per ogni $n \in \mathbf{Z}$, e, perciò, anche $\mu\left(T^n E\right) = 0$ per ogni $n \in \mathbf{Z}$, sicché si avrebbe $\mu_1(E) = 0$, come asserito. Di conseguenza, il problema di trovare la misura ν quando è data μ, equivale a quello di trovare ν quando è data μ_1. D'altro canto, se fosse $\mu_1(E) = 0$, si avrebbe anche $\mu_1\left(TE\right) = \mu_1\left(T^{-1}E\right) = 0$; dunque, si può supporre che T non sia singolare.

Infine, si può limitare la ricerca alle misure invarianti ν che siano equivalenti a μ. Infatti, se esiste una misura invariante σ–finita ν, rispetto alla quale μ è assolutamente continua, esiste un'altra misura invariante e σ–finita ν_0 che è equivalente a μ. Inoltre se ν è finita, tale sarà anche ν_0. Per stabilire questo punto, si ricorra al teorema di RADON–NIKODÝM; esiste una funzione positiva $f \geq 0$ tale che, per ogni $E \in \mathcal{F}$, sia

$$\mu(E) = \int_E f \, d\nu.$$

Posto $A := \{f = 0\}$, si ha $\mu(A) = 0$, e, poiché si è supposto che T non sia singolare, $\mu(T^n A) = 0$ per ogni $n \in \mathbf{Z}$. Si introduca l'insieme

$$B := \cup_{n \in \mathbf{Z}} T^n A$$

che risulta essere misurabile, invariante e di misura μ nulla, $\mu(B) = 0$. Mediante

$$\nu_0(E) := \nu(E \setminus B),$$

si definisca una misura $\nu_0 : \mathcal{F} \longrightarrow \mathbf{R}_+$ che è σ–finita ed invariante rispetto a T. Se $\nu_0(E) = 0$, allora $\nu(E \setminus B) = 0$ e, quindi, $\mu(E \setminus B) = 0$; poiché

$$\mu(E) = \mu(E \cap B) + \mu(E \setminus B),$$

segue che $\mu(E) = 0$. Viceversa, se $\mu(E) = 0$, necessariamente si ha anche $\mu(E \setminus B) = 0$. Se $\omega \in E \setminus B$, allora $f(\omega) > 0$, sicché da

$$\mu(E \setminus B) = \int_{E \setminus B} f \, d\nu$$

segue che $\nu(E \setminus B) = 0$ e, di qui che $\nu_0(E) = 0$.

Il problema della ricerca di una misura invariante è ora ricondotto al seguente: *Data una trasformazione T invertibile, misurabile e non singolare sullo spazio di probabilità $(\Omega, \mathcal{F}, \mu)$ trovare una misura finita (o σ–finita) ν che sia equivalente a μ ed invariante rispetto a T.*

Occorre un nuovo concetto. Sia $\{B_n : n \in \mathbf{N}\}$ una partizione misurabile dell'insieme misurabile A

$$a = \cup_{n \in \mathbf{N}} B_n \qquad \text{e} \qquad E_j \cap E_k = \emptyset \quad (j \neq k).$$

Se $\{C_n : n \in \mathbf{N}\}$ è un'altra successione di insiemi misurabili a due a due disgiunti con unione eguale a C,

$$C == \cup_{n \in \mathbf{N}} C_n \qquad \text{e} \qquad C_j \cap C_k = \emptyset \quad (j \neq k,$$

e se esiste una successione $\{k(n) : n \in \mathbf{N}\}$ di numeri interi tali che per ogni $n \in \mathbf{N}$ sia

$$T^{k(n)} E_n = F_n,$$

allora dice che gli insiemi A e C sono *equivalenti per decomposizione numerabile.*

In analogia con la definizione di DEDEKIND di insieme finito, si dirà che un insieme misurabile A è *limitatto* se A non è equivalente per decomposiozione numerabile ad un suo sottoinsieme proprio. Si dirà che A è σ–limitato se è l'unione di una famiglia numerabile di insiemi limitati. una trasformazione T si dirà limitata (rispettivamente σ–limitata) se l'intero spazio Ω è limitato (rispettivamente σ–limitato).

Se il problema dell'esistenza di una misura invariante ha una soluzione ν, ogni insieme misuraile A che abbia misura finita rispetto a ν, $\nu(A) < +\infty$ è limitato.; se consegue che se ν è una misura finita, T è limitata, se ν è σ–finita, allora T è σ–finita. È stato dimostrato che vale il viceversa: HOPF stabilí che se una trasformazione è limitata, allora esiste una misura finita invariante equivalente a μ, mentre HALMOS dimostrò che se T è σ–limitata, allora esiste una misura invariante σ–finita equivalente a μ.