Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea in Matematica

Una rassegna di teoremi ergodici

Dal teorema ergodico di Birkhoff-Khinchin

al teorema ergodico quantistico di von Neumann

Tesi di Laurea in Matematica

Relatore:
Chiar.mo Prof.
Sandro Graffi

Presentata da:
Roberto Matacchione

Terza Sessione
Anno Accademico 2012-2013

Indice

Introduzione

La teoria ergodica ha le sue origini nella Meccanica statistica. Quest'ultima nacque nell' ottocento ad opera di Helmholtz, Maxwell, Gibbs ed altri, ma soprattutto grazie al genio visionario di Ludwig Boltzmann. Nei suoi lavori della fine dell'ottocento, Boltzmann pose le basi per la spiegazione dei fenomeni termodinamici macroscopici in termini di una teoria atomistica microscopica, cosa che è appunto oggi l'oggetto e la definizione stessa della moderna Meccanica statistica. Ma si deve a Boltzmann, ad esempio, la comprensione del perchè non abbiamo mai osservato un tipico fenomeno macroscopico irreversibile, come un bicchiere d'acqua che cade a terra andando in frantumi, svolgersi 'indietro nel tempo'. Non è, come si credeva prima di Boltzmann, l'irreversibilita del secondo principio della termodinamica a vietare un simile evento, ma semplicemente il fatto che non possiamo osservare l'universo su scale di tempo di gran lunga maggiori della vita dell'universo stesso, perchè un tale fenomeno ha una probabilità non nulla di verificarsi, ma è estremamente piccola, pertanto richiede tempi 'iper-astronomici' per verificarsi. In altre parole, Boltzmann riuscì a conciliare la perfetta reversibilità temporale della meccanica microscopica, come la si legge in tutte le formule, anche della meccanica quantistica, con la quotidiana esperienza della 'freccia del tempo' della termodinamica macroscopica. E l'idea di Boltzmann per trattare la fisica dei sistemi a molti corpi (dell' ordine del numero di Avogadro e anche più) era questa: individuare e contare tutte le configurazioni possibili del sistema. Per fare questo Boltzmann associò, ad ogni sistema di $\mathcal{N}$ particelle (con $\mathcal{N}$ maggiore o dell'ordine di 10^{23}) uno spazio a $6\mathcal{N}$ dimensioni, chiamato da lui spazio delle fasi Γ, tre dimensioni per la posizione e tre dimensioni per le componenti della velocità, per ciascuna delle $\mathcal{N}$ particelle del sistema, il tutto, per semplicità, racchiuso in un contenitore a pareti perfettamente isolanti, di volume $\mathcal{V}$ ed energia costante $\mathcal{E}$. Allora, da un punto di vista puramente matematico, l'evoluzione di questo sistema a molti corpi, sarebbe perfettamente descritta dalla traiettoria del punto di fase nel suo spazio Γ. Il problema, dal punto di vista fisico, è che i nostri cinque sensi, e di coseguenza tutte le nostre apparecchiature strumentali, non sono adeguatamente tarati per osservare su scale spaziali così piccole come quelle atomiche o molecolari. Per di più, la nostra mente e nessun attuale supercalcolatore saprebbe gestire un sistema di 10^{23} equazioni differenziali del moto e, come mostrò Poincarè, occorerebbe anche una conoscenza dei dati iniziali, posizione e velocità, con una precisione di un numero astronomico di cifre decimali. Allora, date questa limitazioni di natura strutturale, Boltzmann ebbe l'idea di racchiudere queste nostre limitazioni dentro quella che chiamò una celletta dello spazio delle fasi. Un volumetto di forma parallelepipeda tanto piccolo quanto occorre per essere compatibile con la massima precisione delle misurazioni possibili sul sistema : all'interno di esso non siamo più in grado di disinguere un punto di fase, cioè un microstato, da un altro. Questa celletta Δ, quindi, contiene tutta l'informazione a noi accessibile sullo stato microscopico del sistema. Ora, se con f indichiamo una qualsiasi grandezza osservabile, che è funzione dello stato microscopico del sistema $f = f(q, v)$ e se l'evoluzione del sistema nello spazio delle fasi Γ la denotiamo genericamente con $\Phi^t(q, v)$, la misurazione sperimentale di tale grandezza si esprime, in forma matematica, come una media temporale,$\overline{f}$, data da

$$\overline{f} = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(\Phi^t(q, v))dt$$

Ma si può anche discretizzare questa formula, passando dal tempo macroscopico,T, a quello molto più piccolo, τ, su cui avvengono i fenomeni microscopici, ponendo $M \stackrel{def}{=} T/\tau$ e riscrivendo il limite precedente come

$$\overline{f} = \lim_{M\to\infty} \frac{f_1 + f_2 + ... + f_M}{M}$$

Essendo incalcolabile anche questa espressione, perchè dipende dalla esatta conoscenza di (q, v) quanto di $\Phi^t(q, v)$ l'idea di Boltzmann fu quella di sostituirla con una media aritmetica. Precisamente, chiamato D il numero di cellette in cui aveva suddiviso lo spazio Γ che corrispondevano allo stesso valore dell'energia totale $\mathcal{E}$ e considerato il valore dell'osservabile f su ognuna di esse, $f(\Delta_i)$, pose

$$\overline{f} = \lim_{M\to\infty} \frac{f_1 + f_2 + ... + f_M}{M} = \frac{1}{D}\sum_{i=1}^{D} f(\Delta_i)$$

Questa fu la formulazione originaria della cosiddetta ipotesi ergodica : poter prevedere il risultato di misurazioni sperimentali per mezzo del calcolo delle probabilità, ma sopratutto, al tempo stesso, giustificarne l'uso.
In questa tesi, si cerca di esporre i principali risultati ottenuti nella teoria ergodica, che è tuttora ben lontana dall'essere una teoria organica e completa.
Nel Cap.1 si forniscono due esempi, la doubling map e le rotazioni di fattore irrazionale sulla circonferenza, dimostrandone l'ergodicità. Si premettono tutti gli strumenti matematici essenziali per arrivare a tale dimostrazione.
Nel Cap.2 si fornisce la dimostrazione del teorema ergodico di Birkhoff-Khinchin, che riguarda le funzioni sommabili, di classe L^1. Questo è in realtà suddiviso in tre teoremi: l'esistenza quasi ovunque della media temporale, l'indipendenza dal punto di fase iniziale, e infine l'uguaglianza con la media in fase nell'ipotesi di indecomponibilità metrica.
Nel Cap.3 si dimostra il cosiddetto teorema ergodico medio di von Neumann, che riguarda le funzioni di classe L^2, premettendo il teorema di ricorrenza di Poincarè, il teorema di Kac e il teorema di Koopman.
Infine, il Cap.4 tratta la dimostrazione del teorema ergodico quantistico di von Neumann, che risale ad un suo articolo del 1929.[Si veda [3]].

Capitolo 1

Sistemi ergodici: definizioni, teoremi ed esempi

1.1 Introduzione: ergodicità

Un sistema dinamico è una coppia (X, T), con X spazio delle fasi e $T\colon X \to X$ una legge che determina come $x \in X$ evolve nel tempo. Se $\Phi^t\colon X \to X$ è un flusso a tempo continuo si può, scegliendo un Δt molto piccolo, considerare la successione $0, \Delta t, 2\Delta t, ..., n\Delta t, ...$, e scrivere, in luogo di $\Phi^{n\Delta t}(x)$ e $\Phi^0(x) = x$, una mappa iterata

$$T^n(x) = (T \circ T \circ ..... \circ T)(x)$$

con $T^0(x) \stackrel{def}{=} x$. Quindi studiamo le iterate della mappa (o trasformazione) $T\colon X \to X$.

Esempio: la doubling map.

$X = [0, 1]$, $\quad T\colon X \to X$ tale che $T(x) \stackrel{def}{=} \begin{cases} 2x & se \quad 0 \le x \le 1/2 \\ 2x - 1 & se \quad 1/2 < x \le 1 \end{cases}$ dove gli estremi 0,1 sono identificati.

Nella doubling map esistono punti cosiddetti periodici. Ad es. è facile verificare che $T^4(2/5) = 2/5$, e si dice che $x = 2/5$ ha periodo 4.

Definizione 1.1.

Un punto $x \in X$ è periodico di periodo $n \in \mathbb{Z}^+$ se n è il più piccolo intero positivo tale che $T^n(x) = x$.

Altri punti della doubling map hanno invece un'orbita in avanti $\mathcal{O}_x^+$, definita come

$$\mathcal{O}_x^+ \stackrel{def}{=} \{x, TxT^2x, ..., T^nx, ...\}$$

che è densa in X.

Definizione 1.2.
$\mathcal{O}_x^+$ è densa in X se per ogni $x' \in X$ e $\epsilon > 0$ esiste un $n_\epsilon \in \mathbb{Z}^+$ tale che $| T^{n_\epsilon}(x) - x' | < \epsilon$.

Si consideri un $[a, b] \subset X$. Ci si può domandare: con che frequenza l'orbita $\mathcal{O}_x^+$ visita l'intervallo $[a, b]$?
Definita $\chi_{[a,b]} \stackrel{def}{=} \begin{cases} 1 & se \quad x \in [a, b] \\ 0 & se \quad x \notin [a, b] \end{cases}$ allora è

$$\sum_{i=0}^{n-1} \chi_{[a,b]}(T^i i(x)) = card\big\{punti\,di\,\mathcal{O}_x^+\,che\,visitano\,[a, b]\,nelle\,prime\,n\,iterate\big\}$$

e quindi

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_{[a,b]}(T^i i(x)) = frazione\,di\,punti\,che\,visitano\,[a, b]\,nelle\,prime\,n\,iterate$$

da cui

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \chi_{[a,b]}(T^i i(x)) = frequenza\,con\,cui\,l'orbita\,di\,x\,giace\,in\,[a, b]$$

Scopo della teoria ergodica è comprendere *quando* tale frequenza coincide con la misura, $b - a$, dell'intervallo $[a, b]$

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_{[a,b]}(T^i i(x)) = b - a = \int_X \chi_{[a,b]} d\mu \quad per\,ogni\,[a, b] \subset X$$

Imporre una tale condizione significa che l'orbita $\mathcal{O}_x^+$ è *equidistribuita* in X e non favorisce una regione di X rispetto ad un'altra. Si può anche andare oltre e sostituire $\chi_{[a,b]}$ con un'arbitraria funzione f e chiedersi:

$$Quando \quad \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) = \int_X f(x)\, d\mu \qquad \forall f \in L^1 \quad ?$$

La risposta a tale quesito è fornita dal Teorema di Birkhoff-Khinchin (che verrà dimostato in seguito, nel Cap.2):

Teorema 1.1.1 (di Birkhoff-Khinchin).
Se $(X, \mathcal{B}, \mu)$ è uno spazio di probabilità e $f \in L^1(X, \mathcal{B}, \mu)$ e se $T: X \to X$ è una trasformazione che preserva la misura ed è ergodica, allora

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int_X f(x)\, d\mu \qquad \mu - quasi\, ovunque\, in\, X.$$

1.2 Un primo esempio di ergodicità: le disribuzioni uniformi modulo 1 di successioni di numeri reali

Sia $(x_n)_{n\in\mathbb{N}}$ una qualsiasi successione di numeri reali e se ne consideri la successione delle parti decimali

$$\{x_n\} \stackrel{def}{=} x_n - [x_n], \qquad [x_n] \stackrel{def}{=} sup\{m \in \mathbb{Z} \mid m \leq x_n\}$$

Poichè è $0 \leq \{x_n\} < 1$, lo studio di $x_n\ mod\, 1$ è lo studio della successione $(\{x_n\})_{n\in\mathbb{N}}$ in $[0, 1]$.

Definizione 1.3.
La successione $(x_n)_{n\in\mathbb{N}}$ è uniformemente distribuita modulo 1 (per brevità udm 1) se e solo se, per definizione,

$$\lim_{n\to\infty} \frac{1}{n} card\{0 \leq j \leq n-1 \mid \{x_j\} \in [a, b]\} = b - a$$

per ogni a e b tali che $0 \leq a < b \leq 1$

Due condizioni necessarie e sufficienti affinchè $(x_n)_{n\in\mathbb{N}}$ sia udm 1 sono date dal seguente

Teorema 1.2.1 (di Weyl).
Sono fatti equivalenti:

1) $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ *è udm* 1
2)per ogni funzione continua $f\colon [0,1] \to \mathbb{R}$ *tale che* $f(0) = f(1)$

$$\lim_{n\to\infty} \sum_{j=0}^{n-1} f(\{x_j\}) = \int_0^1 f(x)dx$$

3)per ogni $l \in \mathbb{Z} \setminus \{0\}$ *si ha*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} e^{2\pi i l x_j} = 0$$

1.2.1 Applicazione del criterio di Weyl

Sia $(x_n)_{n\in\mathbb{N}} = (\alpha n)_{n\in\mathbb{N}} \quad mod\, 1 \quad \alpha \in \mathbb{R}\ fissato$
Risulta che:
per $\alpha \in \mathbb{Q}$ la successione non è udm 1
per $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ la successione è udm 1.
Infatti
se $\alpha \in \mathbb{Q}$) allora $(\alpha\, n \ \ mod\, 1)_{n\in\mathbb{N}}$ possiede solo un numero finito di valori in $[0,1]$ poichè, se $\alpha = p/q\ p,q \in \mathbb{Z},\ \ q \neq 0$ p e q sono coprimi e $p < q$, allora l'orbita è

$$\{0, \frac{p}{q}, \frac{2p}{q}, ..., \frac{(q-1)p}{q}\}$$

essendo $\frac{qp}{q} \equiv 0 \quad mod\, 1$. Ogni elemento ha periodo q.
se $\alpha \in \mathbb{R} \setminus \mathbb{Q}$) si dimostra che $(\alpha\, j \ \ mod\, 1)_{n\in\mathbb{N}}$ è udm 1, per mezzo del criterio di Weyl, come segue.
Posto $x_j \stackrel{def}{=} \alpha\, j \quad j \in \mathbb{N} \quad \alpha \in \mathbb{R} \setminus \mathbb{Q}$ e per $l \in \mathbb{Z} \setminus \{0\}$ (da cui segue che $\alpha\, l$ non è mai intero) si ha:

$$\frac{1}{n} \sum_{j=0}^{n-1} e^{i2\pi l x_j} = \frac{1}{n} \sum_{j=0}^{n-1} e^{i2\pi l \alpha j} = \frac{1}{n} \frac{e^{i2\pi l \alpha n} - 1}{e^{i2\pi l \alpha} - 1}$$

e

$$\mid \frac{1}{n} \frac{e^{i2\pi l \alpha n} - 1}{e^{i2\pi l \alpha} - 1} \mid \leq \frac{1}{n} \frac{2}{\mid e^{i2\pi \alpha l} - 1 \mid} \to 0 \quad n \to \infty$$

1.3 Misure e teorema di estensione di Kolmogoroff

La teoria della misura è lo strumento fondamentale per la teoria ergodica.

1.3.1 Spazi con misura

Una misura è una funzione a valori reali non negativi che, dato un sottoinsieme dello spazio X, dice quanto grande è tale sottoinsieme. Un esempio importante per noi è la misura di Lebesgue su $X = [0, 1]$. In generale, non è possibile definire la misura di un sottoinsieme qualsiasi di $\mathcal{P}(X)$. Occorre restringersi ad una famiglia di sottoinsiemi più piccola di $\mathcal{P}(X)$.

Definizione 1.4.
Una σ-algebra $\mathcal{B}$ è una famiglia di sottoinsiemi di X, $\mathcal{B} \subset \mathcal{P}(X)$, che soddisfa a:
i) $\emptyset \in \mathcal{B}$
ii) se $E \in \mathcal{B}$ allora anche $X \setminus E = \complement E \in \mathcal{B}$
iii) se $\{E_n\}_{n\in\mathcal{N}} \subset \mathcal{B}$ allora anche $\bigcup_{n=1}^{\infty} E_n \in \mathcal{B}$

Definizione 1.5.
Uno spazio misurabile è la coppia $(X, \mathcal{B})$, dove X è un insieme e $\mathcal{B}$ è una σ-algebra su X.

Esempi banali di σ-algebre
1) la σ-algebra $(\emptyset, X)$
2) la σ-algebra $\mathcal{B} = \mathcal{P}(X)$
La prima è troppo piccola, la seconda troppo ampia. Si vedranno in seguito esmpi più interessanti di σ-algebre.

Teorema 1.3.1 (Proprietà delle σ-algebre)**.**
Se $\mathcal{B}$ è una σ-algebra, allora
i) $X \in \mathcal{B}$ (perchè $\complement X = \emptyset \in \mathcal{B}$)
ii) $\bigcap_{n=1}^{\infty} E_n \in \mathcal{B}$ (perchè ogni $\complement E_n \in \mathcal{B}$ e $\complement \bigcup_{n=1}^{\infty} \complement E_n \in \mathcal{B}$).

Definizione 1.6 (σ-algebra di Borel $\mathcal{B}(X)$ su uno spazio metrico compatto X)**.**
E' la σ-algebra minimale che contiene tutti gli aperti di X.

Si può costruire la σ-algebra di Borel sui reali $\mathcal{B}(\mathbb{R})$ in diversi modi equivalenti. Può essere infatti generata, ad esempio, da:
1) gli intervalli chiusi $[a, b]$, con $a < b$
2) gli intervalli aperti (a, b), con $a < b$
3) le semirette sinistre chiuse $(-\infty, a]$
4) le semirette destre aperte$(a, +\infty)$
5) gli insiemi aperti di $\mathbb{R}$.

Definizione 1.7.
La funzione $\mu\colon \mathcal{B} \to \mathbb{R}^+$ è una misura finita se:
i) $\mu(\emptyset) = 0$
ii) se $E_n \cap E_m = \emptyset$ per $n \neq m$ e $\bigcup_{j=1}^{\infty} E_j \in \mathcal{B}$, allora $\mu(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu(E_j)$

Chiamiamo la terna $(X, \mathcal{B}, \mu)$ uno spazio con misura e, se $\mu(X) = 1$, uno spazio di probabilità.

Definizione 1.8.
Una proprietà vale quasi ovunque (q.o.) o μ-quasi ovunque (μ-q.o.) se vale ad eccezione di al più un insieme di misura nulla.

Ad esempio, i punti del degmento $[0, 1]$, secondo la misura μ di Lebesgue, sono quasi ovunque irrazionali perchè i punti razionali costituiscono un insieme di misura nulla.

Definizione 1.9.
Un'algebra $\mathcal{A}$ di sottoinsiemi di X è una famiglia $\mathcal{A} \subset \mathcal{P}(X)$ che soddisfa:
i) $\emptyset \in \mathcal{A}$
ii) se $A \in \mathcal{A}$ allora anche $\complement A \in \mathcal{A}$
iii) se $A_1, A_2, ..., A_n \in \mathcal{A}$ allora $\bigcup_{j=1}^{n} A_j \in \mathcal{A}$

1.3.2 Il teorema di estensione di Kolmogoroff

Teorema 1.3.2 (di estensione di Kolmogoroff)**.**
Dati un'algebra $\mathcal{A}$ di sottoinsiemi di X ed una misura $\mu\colon \mathcal{A} \to [0, 1]$ tale che
i) $\mu(\emptyset) = 0$
ii) se $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, $A_i \neq A_j$ per $i \neq j$ allora $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

iii) $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$
allora
esiste un'unica misura $\mu' \colon \mathcal{B}(\mathcal{A}) \to [0,1]$ *che è estensione di* μ *da* $\mathcal{A}$ *a* $\mathcal{B}(\mathcal{A})$.

Applicazioni del teorema di estensione
Si userà il teorema di estensione nei seguenti modi.
1)Prenderemo $X = [0,1]$ e $\mathcal{A}$ l'algebra di tutte le unioni finite di sottointervalli di X. Definiremo una misura μ dei sottointervalli in modo consistente con le ipotesi del teorema di Kolmogoroff. Seguirà allore che μ definisce, per estensione, una misura sulla σ-algebra di Borel.
2) Per mostrare che due σ-algebre di Borel con misura su $[0,1]$ danno la stessa misura, sarà sufficiente mostrare che esse forniscono la stessa misura su ogni sottointervallo.
3) Useremo il teorema di estensione anche per mostrare la T-invarianza di una misura.

1.4 Integrazione di Lebesgue.

Sia $(X, \mathcal{B}, \mu)$ uno spazio con misura. Vogliamo integrare le funzioni $f \colon X \to \mathbb{R}, \mathbb{C}$ rispetto alla misura μ.

Definizione 1.10.
Una funzione $f \colon X \to \mathbb{R}$ è misurabile se e solo se, per definizione, $f^{-1}(D) \in \mathcal{B}$ per ogni boreliano $D \in \mathcal{B}$.
Una funzione $f \colon X \to \mathbb{C}$ è misurabile se e solo se, per definizione, lo sono $Re\{f\}$ e $Im\{f\}$.

Questa definizione è equivalente al richiedere che $f^{-1}(-\infty, c) \in \mathcal{B}$ per ogni $c \in \mathcal{R}$

Definizione 1.11.
Una funzione semplice f_s è una combinazione lineare di funzioni caratteristiche χ_{B_j} di sottoinsiemi B_j due a due disgiunti:

$$f_s \stackrel{def}{=} \sum_{j=1}^{r} a_j \chi_{B_j} \quad a_j \in \mathcal{R} \quad B_k \cap B_l = \emptyset \ \textit{per} \ k \neq l$$

Definizione 1.12. L'integrale di Lebesgue di una funzione semplice è, per definizione:

$$\int_X f_s d\mu \overset{def}{=} \sum_{j=1}^{r} a_j \mu(B_j)$$

1.4.1 Costruzione dell'integrale di Lebesgue mediante funzioni semplici.

Se $f\colon X \to \mathbb{R}$ è misurabile ed è $f \geq 0$, esiste una successione crescente di funzioni semplici che converge a f quando $n \to \infty$, $f_{s_n} \nearrow f$ e si pone

$$\int_X f\, d\mu \overset{def}{=} \lim_{n\to\infty} \int_X f_{s_n}\, d\mu$$

Tale integrale esiste (potendo anche essere $+\infty$), ma non dipende dalla scelta della successione $\big\{f_{s_n}\big\}_{n\in\mathbb{N}}$ Per un'arbitraria funzione $f\colon \ \to \mathbb{R}$ si pone $f \overset{def}{=} f^+ - f^-$, dove

$$f^+ \overset{def}{=} \max\big\{f, 0\big\} \geq 0 \quad e \quad f^- \overset{def}{=} \max\big\{-f, 0\big\} \geq 0$$

e quindi si definisce

$$\int_X f\, d\mu \overset{def}{=} \int_X f^+ d\mu - \int_X f^- d\mu$$

Per $f\colon X \to \mathbb{C}$ si pone $\int_X f\, d\mu \overset{def}{=} \int_X Re\big\{f\big\} + i \int_X Im\big\{f\big\}$.

Definizione 1.13.
f è sommabile se e solo se, per definizione, $\int_X |\ f\ |\ d\mu < +\infty$.
Lo spazio delle funzioni $f\colon X \to \mathbb{C}$, sommabili, è denotato $L^1(X, \mathcal{B}, \mu)$.

Definizione 1.14.
Sia $(X, \mathcal{B}, \mu)$ uno spazio di probabilità (cioè t.c. $\mu(X) = 1$) e sia $B \in \mathcal{B}$. Alllora $\chi_B f \in L^1$ e si definisce

$$\int_B f d\mu \overset{def}{=} \int_X \chi_B f d\mu$$

Definizione 1.15. Due funzioni misurabili $f, g\colon X \to \mathbb{R}, \mathbb{C}$ sono dette equivalenti o uguali q.o. se $f = g \quad \mu - q.o.$ cioè se $\mu\big\{x \in X \mid f(x) \neq g(x)\big\} = 0$

Lemma 1.4.1.
Se $f, g \in L^1(X, \mathcal{B}, \mu)$ *e se* $f = g$ *q.o. allora* $\int_X f d\mu = \int_X g d\mu$

Ne consegue che $L^1(X, \mathcal{B}, \mu)$ è in realtà un insieme di classi di equivalenza di funzioni $f, g \colon X \to \mathbb{C}$ sommabili e tali che $f = g$ *q.o.*

Definizione 1.16.
La norma in L^1 è $\|f\|_1 \overset{def}{=} \int_X |\, f \,|\, d\mu$
La metrica di L^1 è $d(f,g) \overset{def}{=} \|f - g\|_1$

Con tali norma e metrica L^1 risulta essere uno spazio vettoriale normato e completo, cioè uno spazio di Banach.

Definizione 1.17.
Lo spazio $L^2(X, \mathcal{B}, \mu)$ consiste delle classi di equivalenza delle funzioni $f, g \colon X \to \mathbb{C}$ tali che $\int_X |\, f \,|^2 d\mu < +\infty$ con metrica data da $d(f,g) \overset{def}{=} \|f - g\|_2 \overset{def}{=} \left(\int_X |\, f - g \,|^2 d\mu \right)^{1/2}$

Osservazione 1.
In uno spazio di probabilità ($\mu(X) = 1$) si ha $L^2 \subset L^1$ perchè, per la disuguaglianza di Cauchy-Schwarz, si ha, posto $g(x) = 1(x)$

$$\int_X |\, f \,|\, d\mu \leq \left(\int_X |\, f \,|^2 d\mu \right)^{1/2} \left(\mu(X) \right)^{1/2} < +\infty$$

Teorema 1.4.2 (della convergenza dominata)**.**
Sia $g \colon X \to \mathbb{R}$ *sommabile e* $f_n \colon X \to \mathbb{R}$ *una successione di funzioni misurabili tali che* $|\, f_n \,| \leq g$ *q.o. e* $\lim_{n \to +\infty} f_n = f$ *q.o.*
Allora anche f *è sommabile e si ha*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \lim_{n\to\infty} f_n d\mu = \int_X f d\mu$$

Teorema 1.4.3 (della convergenza monotona)**.**
Se $f_n \colon X \to \mathbb{R}$ *è una successione crescente di funzioni sommabili tali che* $\exists M > 0$ *per cui* $|\int_X f_n d\mu| < M$ *per ogni* n *allora* $f(x) = \lim_{n\to\infty} f_n(x)$ *esiste* $\mu-$*q.o. e si ha*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X lim_{n\to\infty} f_n d\mu = \int_X f d\mu$$

1.5 Misure invarianti

Siano $(X, \mathcal{B}, \mu)$ uno spazio di probabilità e (X, T) un sistema dinamico. Se $B \in \mathcal{B}$ definiamo $T^{-1}(B) \stackrel{def}{=} \{x \in X \mid Tx \in B\}$. Non assumiamo che T sia una biiezione.

Definizione 1.18.
La trasformazione $T\colon X \to X$ è misurabile se per ogni $B \in \mathcal{B}$ si ha $T^{-1}(B) \in \mathcal{B}$.

Osservazione 2.
Poichè lavoreremo con X spazio metrico compatto dotato di σ-algebra di Borel, ogni trasformazione continua è pure misurabile (perchè la retroimmagine di un aperto è un aperto).

Osservazione 3.
Sia $\mathcal{A}$ un'algebra che genera la σ-algebra di Borel $\mathcal{B}(\mathcal{A})$. Allora se per ogni $A \in \mathcal{A}$ è $T^{-1}(A) \in \mathcal{B}$, T è misurabile.

Definizione 1.19.
μ è misura T-invariante ovvero T preserva la misure μ, se per ogni $B \in \mathcal{B}$ si ha che $T^{-1}(B) \in \mathcal{B}$

1.5.1 Uso del teorema di estensione di Kolmogoroff per dimostrare l'invarianza

Useremo il teorema di estensione per provare l'invarianza della misura di Lebesgue μ in due casi, servendoci anche del seguente

Corollario 1.5.1 (Corollario del teorema di estensione di Kolmogoroff)**.**
Se μ_1 e μ_2 sono due misure su $\mathcal{B}(\mathcal{A})$ tali che $\mu_1(A) = \mu_2()A$ per tutti gli $A \in \mathcal{A}$, allora $\mu_1 = \mu_2$ su $\mathcal{B}(\mathcal{A})$

Definizione 1.20 (misura $T_*\mu$)**.**
Siano $(X, \mathcal{B}, \mu)$ uno spazio di probabilità e $T\colon X \to X$ misurabile. Per ogni $B \in \mathcal{B}$ si definisce
$T_*\mu(B) \stackrel{def}{=} \mu(T^{-1}(B))$

$T_*\mu$ è una misura di probabilità su $(X, \mathcal{B}, \mu)$. μ è misura T-invariante se e solo se $T_*\mu = \mu$.
Il corollario precedente stabilisce che se due misure coincidono sull'algebra $\mathcal{A}$ allora esse coincidono sulla σ-algebra $\mathcal{B}(\mathcal{A})$ generata da $\mathcal{A}$.
Quindi, se si può mostrare che $T_*\mu = \mu$ su $\mathcal{A}$, che genera $\mathcal{B}(\mathcal{A})$, resta dimostrato che $T_*\mu = \mu$ su $\mathcal{B}(\mathcal{A})$.
Questo è quello che faremo nei prossimi due esempi.

Esempi: la doubling map e le rotazioni sulla circonferenza $\mathbb{R}/\mathbb{Z}$

Esempio 1. La doubling map
Sia $X = [0, 1]$ ad estremi identificati, $\mathcal{B}$ la σ-algebra di Borel, μ la misura di Lebesgue. La trasformazione T sia definita da $\quad T(x) \overset{def}{=} 2x \quad mod\, 1$.

Teorema 1.5.2 (T-invarianza della misura di Lebesgue nella doubling map)**.**

Se T è la doubling map, la misura di Lebesgue è T-invariante.

Dimostrazione.
Come conseguenza del teorema di estensione, è sufficiente mostrare che la tesi vale sulle unioni finite di sottointervalli che costituiscono l'algebra $\mathcal{A}$, che genera la σ-algebra $\mathcal{B}$ di Borel.
Per un singolo intervallo $[a, b]$ con $0 \leq a < b \leq 1$, si ha

$$T^{-1}([a,b]) = \Big\{ x \in \mathbb{R}/\mathbb{Z} \mid T(x) \in [a,b] \Big\} = \left[\frac{a}{2}, \frac{b}{2}\right] \cup \left[\frac{a+1}{2}, \frac{b+1}{2}\right] \tag{1.1}$$

da cui

$$\begin{aligned} T_*\mu([a,b]) == \mu\big(T^{-1}([a,b])\big) &= \mu\Big([\frac{a}{2}, \frac{b}{2}] \cup [\frac{a+1}{2}, \frac{b+1}{2}] \Big) = \\ &= \frac{b}{2} - \frac{a}{2} + \frac{b+1}{2} - \frac{a+1}{2} = b - a = \mu([a,b]) \end{aligned} \tag{1.2}$$

Quindi $T_*\mu = \mu$ su $\mathcal{A}$ e, per il corollario del teorema di estensione, segue

$$T_*\mu = \mu \quad su \quad \mathcal{B}(\mathcal{A}) \tag{1.3}$$

Si conclude che la misura di Lebesgue è T-invariante, come dovevasi dimostrare. $\square$

Esempio 2. Le rotazioni sulla circonferenza $\mathbb{R}/\mathbb{Z}$
Sia $T\colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ tale che $x \mapsto T(x) \stackrel{def}{=} x + \alpha \quad mod\, 1$, $\mathcal{B}$ la σ-algebra di Borel e μ la misura di Lebesgue.

Teorema 1.5.3 (Invarianza della misura di Lebesgue nelle rotazioni sulla circonferenza)**.**
Per la T definita sopra, la misura di Lebesgue μ è T-invariante

Dimostrazione.
Usiamo ancora una volta il teorema di estensione. Ora

$$T^{-1}([a,b]) = [a-\alpha, b-\alpha] \quad mod\ 1 \tag{1.4}$$

da cui

$$T_*\mu([a,b]) = \mu(T^{-1}([a,b])) = \mu([a-\alpha, b-\alpha]) = b-a = \mu([a,b]) \tag{1.5}$$

e se $\quad a-\alpha < 0 < b-\alpha \quad$ allora

$$T^{-1}([a,b]) = [0, b-\alpha] \cup [a-\alpha+1, 1] \tag{1.6}$$

e anche in questo caso

$$\mu(T^{-1}([a,b])) = b-\alpha+1-(a-\alpha+1) = b-a = \mu([a,b]) \tag{1.7}$$

La conclusione è che la misura di Lebesgue μ è T-invariante su tutta la σ-algebra $\mathcal{B}(\mathcal{A})$, come dovevasi dimostrare. □

1.6 Misure ergodiche

Definizione 1.21 (Ergodicità, o metrica indecomponibilità)**.**
Sia $(X, \mathcal{B}, \mu)$ uno spazio di probabilità, $T\colon X \to X$ tale che $\mu(T^{-1}(B)) = \mu(B)$ per ogni $B \in \mathcal{B}$.
T è ergodica rispetto a μ, ovvero μ è una misura ergodica, se e solo se, per definizione, ogniqualvolta $B \in \mathcal{B}$ soddisfa a $T^{-1}(B) = B$ allora è o $\mu(B) = 0$ oppure $\mu(B) = 1$.

Osservazione 4 (Metrica indecomponibilità).
L'ergodicità di μ può essere vista come una condizione di indecomponibilità metrica nel senso seguente. Se l'ergodicità non vale, si può trovare un boreliano $B \in \mathcal{B}$ tale che $\mu(T^{-1}(B)) = \mu(B)$ e $0 < \mu(B) < 1$.
In tal caso si può decomporre $T\colon X \to X$ in $T\colon B \to B$ con misura invariante $\frac{1}{\mu(B)}\mu(. \cap B)$ e
$T\colon X \setminus B \to X \setminus B$ con misura invariante $\frac{1}{1-\mu(B)}\mu(. \cap (X \setminus B))$.

Teorema 1.6.1 (Caratterizzazioni dell'ergodicità)**.**
Sia T tale che $\mu(T^{-1}B) = \mu(B)$ sullo spazio di probabilità $(X, \mathcal{B}, \mu)$. Sono fatti equivalenti:
i) T è ergodica
ii) ogniqualvolta $f \in L^1(X, \mathcal{B}, \mu)$ soddisfa a $f \circ T = f$ q.o. allora $f = cost.$ q.o.
iii)ogniqualvolta $f \in L^2(X, \mathcal{B}, \mu)$ soddisfa a $f \circ T = f$ q.o. allora $f = cost.$ q.o.

Osservazione 5.
Se f è costante allora ovviamente $f \circ T = f$. Quindi la proposizione afferma che *le uniche* funzioni T-invarianti sono le funzioni *costanti*, se T è ergodica.

Dimostrazione.
i) $\Rightarrow$ ii)
Per $k \in \mathbb{Z},\ n \in \mathbb{N}$, si costruiscano i boreliani

$$X(k,n) \overset{def}{=} \left\{ x \in X \mid \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \right\} = f^{-1}\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \tag{1.8}$$

Poichè f è misurabile $X(k,n) \in \mathcal{B}$. Si ha allora

$$T^{-1}(X(k,n)) \,\Delta\, X(k,n) \subset \left\{ x \in X \mid f(T(x)) \neq f(x) \right\} \tag{1.9}$$

(perchè $A \Delta B$ non contiene $A \cap B$), cosicchè essendo per ipotesi $f \circ T = f \quad q.o.$, si ha

$$\mu\Big(T^{-1}(X(k,n)) \,\Delta\, X(k,n) \Big) = 0 \tag{1.10}$$

da cui essendo per ipotesi T ergodica deve essere $\mu(X(k,n)) = 0$ oppure 1.
Essendo f sommabile deve anche essere finita q.o., da cui per ogni n fissato

$$f^{-1}(\mathbb{R}) = f^{-1}\left(\bigcup_{k=-\infty}^{+\infty} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right) = \bigcup_{k=-\infty}^{+\infty} f^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right) = \bigcup_{k=-\infty}^{+\infty} X(k,n) = X \quad q.o. \tag{1.11}$$

e quindi deve essere

$$\mu\Big(X\,\Delta \bigcup_{k=-\infty}^{+\infty} X(k,n)\Big) = 0 \tag{1.12}$$

Dunque si ha

$$\sum_{k=-\infty}^{+\infty} \mu(X(k,n)) = \mu(X) = 1 \tag{1.13}$$

da cui segue necessariamente che deve esistere un unico k_n per il quale è

$$\mu(X(k_n,n)) = 1 \tag{1.14}$$

Allora, costruendo l'insieme $Y \stackrel{def}{=} \bigcap_{n=1}^{\infty} X(k_n,n)$, si ha in primo luogo che $\mu(Y) = 1$ e, in secondo luogo, per ogni $x, y \in Y$ $f(x), f(y) \in \left[\frac{k_n}{2^n}, \frac{k_n+1}{2^n}\right)$ per tutti gli $n \geq 1$ e quindi $|\,f(x) - f(y)\,| \leq \frac{1}{2^n}$ per tutti gli $n \geq 1$. Segue allora che $f(x) = f(y)$ e cioè f è costante su Y. Ma allora f è costante anche su X perchè $\mu(Y) = 1 = \mu(X)$, come doveva mostrarsi.
ii)⇒iii)
Segue dal fatto che $f \in L^2 \Rightarrow f \in L^1$.
iii)⇒i)
Supponiamo che $B \in \mathcal{B}$ sia tale che $T^{-1}B = B$. Allora è $\chi_B \in L^2(X, \mathcal{B}, \mu)$ e si ha

$$(\chi_B \circ T)(x) = (\chi_{T^{-1}B})(x) = \chi_B(x)$$

per l'ipotesi $T^{-1}B = B$ e per ogni $x \in X$. Quindi χ_B è, per ipotesi, costante q.o. in X e potendo essere solo $\chi_B(x) = 1 \quad q.o.$ oppure $\chi_B(x) = 0 \quad q.o.$ si ha

$$\int_X \chi_B(x)\,d\mu = \mu(B) = \begin{cases} 0 \qquad \mu - q.o. \\ \qquad oppure \\ 1 \qquad \mu - q.o \end{cases}$$

ovvero T è ergodica rispetto a μ, perchè $T^{-1}B = B$ implica $\mu(B) = 0$ oppure 1.
Il teorema è così dimostrato. □

1.7 Uso delle serie di Fourier per dimostrare l'ergodicità

Useremo un metodo basato sulle serie di Fourier per dimostrare l'ergodicità della doubling map e delle rotazioni sulla circonferenza $\mathbb{R}/\mathbb{Z}$ di un fattore α irrazionale.

1.7.1 Serie di Fourier

Associamo ora, ad ogni funzione $f \in L^2(X, \mathcal{B}, \mu)$, la sua serie di Fourier $\sum_{n=-\infty}^{+\infty} c_n(f)e^{2\pi inx}$ con

$$c_n(f) \stackrel{def}{=} < f, e_n > = \int_0^1 f(x)e^{-2\pi inx}d\mu$$

dove si è usato il sistema ortonormale completo $\{e_n\}_{n\in\mathbb{Z}}$ dato da $e_n(x) \stackrel{def}{=} e^{2\pi inx}$, avendo posto

$$< f, g > \stackrel{def}{=} \int_X f\overline{g}d\mu \quad e \quad d(f, g) \stackrel{def}{=} \left(\int_X |\, f - g\,|^2 \, d\mu\right)^{1/2}$$

Vale la seguente

Proposizione 1.7.1.
1) Siano $f, g \in L^2(X, \mathcal{B}, \mu)$. Allora $f = g \quad \mu - q.o.$ se e solo se $c_n(f) = c_n(g) \quad \forall n \in \mathbb{Z}$.
2) (Teorema di Riemann-Lebesgue)
Se $f \in L^2$ allora $c_n(f) \to 0$ per $|\, n\,| \to \infty$

Faremo uso anche del seguente

Teorema 1.7.2 (Caratterizzazioni di T che preserva μ)**.**
Sia $T : X \to X$ una trasformazione misurabile in uno spazio di probabilità $(X, \mathcal{B}, \mu)$. Sono equivalenti:
i) T preserva μ: $\mu(T^{-1}B) = B$
ii) per ogni $f \in L^1$ si ha $\int_X f d\mu = \int_X f \circ T d\mu$
iii) per ogni $f \in L^2$ si ha $\int_X f d\mu = \int_X f \circ T d\mu$

Dimostrazione.
Si premette il seguente

Lemma 1.7.3.
$\chi_{T^{-1}B} = \chi_B \circ T$

Dimostrazione.
$\int \chi_{T^{-1}B}(x)d\mu = \int \chi_B(T(x))d\mu = \int(\chi_B \circ T)(x)d\mu$ □

i) $\Rightarrow$ ii)
Per ogni funzione caratteristica $\chi_B,\ B \in \mathcal{B}$, si ha

$$\int \chi_B d\mu = \mu(B) = \mu(T^{-1}B) = \int \chi_{T^{-1}B} d\mu = \int (\chi_B \circ T) d\mu$$

e queste uguaglianze valgono anche per ogni funzione semplice. Si è visto che, data una qualsiasi funzione $f \in L^1$, $f \geq 0$, si può trovare una successione crescente di funzioni semplici $\left\{f_{s_n}\right\}_{n\in\mathbb{N}}$ e per ogni n si ha

$$\int f_{s_n} d\mu = \int f_{s_n} \circ T\ d\mu$$

Allora, applicando il teorema della convergenza monotona ad entrambi i membri, si ottiene

$$\int f d\mu = \int f \circ T\ d\mu$$

ii)$\Rightarrow$iii)
Segue dal fatto che, per uno spazio di probabilità,vale $L^2(X, \mathcal{B}, \mu) \subset L^1(X, \mathcal{B}, \mu)$.
iii)$\Rightarrow$i)
Sia $B \in \mathcal{B}$. Allora $\chi_B \in L^2$, e si ha

$$\int |\ \chi_B\,|^2\, d\mu = \int \chi_B d\mu = \mu(B) = \int (\chi_B \circ T) d\mu = \int \chi_{T^{-1}B} d\mu = \mu(T^{-1}B)$$

Il teorema è dimostrato. □

Se ora associamo ad $f \in L^2$ la sua serie di Fourier e ne consideriamo la ridotta n-esima

$$s_n(x) \overset{def}{=} \sum_{l=-n}^{n} c_l(f) e^{2\pi i l x} \tag{1.15}$$

abbiamo la convergenza in media quadratica

$$\|s_n(x) - f(x)\|_2 \to 0 \qquad n \to \infty \tag{1.16}$$

Poichè T preserva la misura μ, usando il teorema precedente, si ha

$$\|s_n \circ T - f \circ T\|_2 = \left(\int | s_n \circ T - f \circ T |^2 d\mu \right)^{1/2} = \left(\int | s_n - f |^2 \circ T d\mu \right)^{1/2} =$$
$$= \left(\int | s_n - f |^2 d\mu \right)^{1/2} = \|s_n - f\|_2 \to 0 \qquad n \to \infty \tag{1.17}$$

Il metodo che ora siamo in grado di applcare, per dimostrare l'ergodicità, consiste in quanto segue.
Per l'unicità della serie di Fourier, se $S_n \circ T$ è una serie di Fourier essa deve essere la serie di Fourier di $f \circ T$. Se prendiamo la serie di Fourier di f valutata su $T(x)$, otteniamo la serie di Fourier di $(f \circ T)(x)$.
Ma, se T è ergodica e se è $f \circ T = f \quad \mu - q.o.$ si possono paragonare i coefficienti di Fourier, $c_n(f)$ *e*
$c_n(f \circ T)$ di $f(x)$ *e* $f(T(x))$ rispettivamente, per ottenere delle relazioni tra questi coefficienti e quindi verificare se si devono annullare tutti ad eccezione di c_0. Se ciò deve necessariamente verificarsi, cioè se è $f = c_0 = cost.\ \mu - q.o.$, T è ergodica.

1.8 Dimostrazioni di ergodicità

Siamo ora in grado di dimostrare l'ergodicità delle rotazioni sulla circonferenza $\mathbb{R}/\mathbb{Z}$, di un fattore $\alpha \in \mathbb{R}/\mathbb{Z}$, e della duobling map.

1.8.1 Le rotazioni sulla circonferenza $\mathbb{R}/\mathbb{Z}$

Sia $T(x) \stackrel{def}{=} x + \alpha \quad mod\ 1$ su $\mathbb{R}/\mathbb{Z}$.
i) se $\alpha \in \mathbb{Q}$, T non è ergodica rispetto alla misura di Lebesgue
ii) se $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, T è ergodica.
Infatti
i) se $\alpha \in \mathbb{Q}$ sia $\alpha = \frac{p}{q}$ con $p, q \in \mathbb{Z}$ coprimi $p < q$ *e* $q \neq 0$. Si definisca

$f(x) \stackrel{def}{=} e^{2\pi iqx} \in L^2$. Allora si ha $f(T(x)) = e^{2\pi iq(x+p/q)} = e^{2\pi i(qx+p)} = e^{2\pi iqx} = f(x) \neq cost.$
Quindi T non è ergodica.
ii) Sia $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Si supponga che $f \in L^2$ e che $f \circ T = f \quad \mu - q.o.$. Si ponga

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n(f) e^{2\pi inx}$$

Allora

$$f(T(x)) = \sum_{n=-\infty}^{+\infty} c_n(f) e^{2\pi inx} e^{2\pi in\alpha}$$

Paragonando i coefficienti si ha

$$c_n(f) = c_n(f) e^{2\pi in\alpha}$$

che è possibile, per α irrazionale, solo se $c_n(f) = 0$ per ogni n intero diverso da 0. Segue che è $f(x) = c_0 = cost.$ e T è ergodica.

1.8.2 La doubling map

Sia $T(x) \stackrel{def}{=} 2x \quad mod\ 1$. Sia $f \in L^2$ e $f \circ T = f \quad \mu - q.o.$ Sia $f(x) = \sum_{n=-\infty}^{+\infty} c_n(f) e^{2\pi inx}$.
Per ogni $p > 0$ intero, $f \circ T^p$, ha sviluppo in serie di Fourier

$$f(T^p(x)) \sum_{n=-\infty}^{+\infty} c_n(f) e^{2\pi in2^px} = \sum_{n=-\infty}^{+\infty} c_{n2^p}(f) e^{2\pi inx}$$

Paragonando i due sviluppi si ottiene

$$c_n = c_{n2^p} \quad per\ ogni\ n \in \mathbb{Z} \quad e\ per\ ogni\ p = 1, 2, ...$$

Se fissiamo un $\overline{n} \neq 0$, si ha $\overline{n}2^p \to \infty \quad$ per $p \to \infty$.
Ma, per il teorema di Riemann-Lebesgue, deve essere $c_{\overline{n}2^p} \to \infty$ per $p \to \infty$.
Ciò è possibile solo se $c_n = 0$ per ogni $n \neq 0$.
Dunque $f = c_0 = cost.\ \mu - q.o.$ e si conclude che T è ergodica.

Capitolo 2

Il teorema ergodico di Birkhoff-Khinchin

2.1 Il teorema di Liouville

Sia dato il sistema di equazioni di Hamilton

$$\begin{cases} \dfrac{dq_i}{dt} = \dfrac{\partial H}{\partial p_i} \\ \dfrac{dp_i}{dt} = -\dfrac{\partial H}{\partial q_i} \end{cases} \quad per \quad i = 1, ..., n$$

Il valore di$(q_1(0), p_1(0), ..., q_n(0), p_n(0))$ all'istante $t = t_0$ determina univocamente il valore di
$(q_1(t), p_1(t), ..., q_n(t), p_n(t))$ per ogni $t > t_0$, poichè le equazioni del sistema canonico sono del primo ordine.
Si può quindi definire un'applicazione di avanzamento $\Phi^t : \Gamma \longrightarrow \Gamma$ tale che $q(0) \mapsto q_(t) = \Phi^t(q(0))$.
Liouville scoprì che Φ^t conserva lamisura di un qualunqe sottoinsieme $M \subset \Gamma$, cioè che il flusso nello spazio delle fasi Γ è simile a quello di un fluido incomprimibile.

Teorema 2.1.1 (Teorema di Liouville).

Sia M un sottoinsieme misurabile secondo Lebesgue di Γ, $M \subset \Gamma$, di misura $\mu(M)$. Durante il moto M viene portato in $\Phi^t M$.
Allora, per ogni t, $\mu(\Phi^t M) = \mu(M)$.

Dimostrazione.
Per maggiore chiarezza e semplicità, si effettui il seguente cambiamento di variabili:

$$x_i \stackrel{def}{=} q_i \qquad x_{N+i} \stackrel{def}{=} p_i \qquad i = 1, ..., N \tag{2.1}$$

$$X_i \stackrel{def}{=} \frac{\partial H}{\partial q_i} \qquad X_{N+i} \stackrel{def}{=} -\frac{\partial H}{\partial q_i} \qquad i = 1, ..., N \tag{2.2}$$

Con queste notazioni il sistema canonico diventa

$$\frac{dx_i}{dt} = X_i(x_1, x_2, ..., x_{2N}) \qquad i = 1, 2, ..., 2N \tag{2.3}$$

Con condizioni iniziali $(x_1^0, x_2^0, ..., x_{2N}^0)$ questo sistema ha un'unica ben determinata soluzione data da

$$\begin{aligned} x_1 &= f_1(t; x_1^0, ..., x_{2N}^0) \\ &\vdots \\ x_{2N} &= f_{2N}(t; x_1^0, ..., x_{2N}^0) \end{aligned} \tag{2.4}$$

Sia $\mu(M)$ la misura di un sottoinsieme $M \subset \Gamma$. Allora

$$\mu(\Phi^t M) = \int_{\Phi^t M} dx_1, ..., dx_{2N} \tag{2.5}$$

In questo integrale, si effettui il cambiamento di variabili

$$x_i \stackrel{def}{=} f_i(t; y_1, ..., y_{2N}) \qquad per\ i = 1, 2, ..., 2N \tag{2.6}$$

(dove t è un parametro).
Allora il punto $(y_1, ..., y_{2N}) \in \Gamma$ descrive l'insieme M quando il punto $(x_1, ..., x_{2N})$ descrive l'insieme $\Phi^t M$. E quindi

$$\mu(\Phi^t M) = \int_M J(t; y_1, ...y_{2N})\, dy_1.....dy_{2N} \tag{2.7}$$

dove

$$J = J(t; y_1, ..., y_{2N}) \stackrel{def}{=} det\frac{\partial(x_1, ..., x_{2N})}{\partial(y_1, ..., y_{2N})} \tag{2.8}$$

Derivando rispetto a t

$$\frac{d}{dt}\mu(\Phi^t M) = \int_M \frac{\partial J}{\partial t}\, dy_1.....dy_{2N} \tag{2.9}$$

Si calcola ora la derivata dello jacobiano mediante la regola di derivazione dei determinanti, ovvero

$$\frac{\partial J}{\partial t} = \sum_{i=1}^{2N} J_i \quad con \quad J_i \stackrel{def}{=} \frac{\partial(x_1, ..., \frac{\partial x_i}{\partial t}, ..., x_{2N})}{\partial(y_1, ..., y_{2N})} \qquad per \quad i = 1, ..., 2N \tag{2.10}$$

In base al sistema canonico, è $\frac{\partial x_i}{\partial t} = \frac{dx_i}{dt}$ e dunque

$$J_i = \frac{\partial(x_1, ..., X_i, ..., x_{2N})}{\partial(y_1, ..., y_{2N})} \quad i = 1, ..., 2N \tag{2.11}$$

Ma $\dfrac{\partial X_i}{\partial y_k} = \displaystyle\sum_{r=1}^{2N} \frac{\partial X_i}{\partial x_r}\frac{\partial x_r}{\partial y_k}$, $1 \leq i, k \leq 2N$, da cui $J_i = \displaystyle\sum_{r=1}^{2N} \frac{\partial X_i}{\partial x_r}\frac{\partial(x_1, ..., x_{i-1}, x_r, x_{i+1}, ..., x_{2N})}{\partial(y_1, ..., y_{2N})}$
Ora, chiaramente

$$\frac{\partial(x_1, ..., x_{i-1}, x_r, x_{i+1}, ..., x_{2N})}{\partial(y_1, ..., y_{2N})} = \begin{cases} J & se \quad r = i \\ 0 & se \quad r \neq i \end{cases} \tag{2.12}$$

donde $J_i = \frac{\partial X_i}{\partial x_i} J$.

Sostituendo nella prima delle (10) e notando che è $\displaystyle\sum_{i=1}^{2N} \frac{\partial X_i}{\partial x_i} = \sum_{i=1}^{2N} \frac{\partial^2 H}{\partial p_i \partial q_i} - \sum_{i=1}^{2N} \frac{\partial^2 H}{\partial q_i \partial p_i} =$ 0 si ottiene
$\dfrac{\partial J}{\partial t} = \displaystyle\sum_{i=1}^{2N} J_i = J \sum_{i=1}^{2} N \frac{\partial X_i}{\partial x_i} = 0.$
La conclusione è

$$\frac{d}{dt}\mu(\Phi^t M) = 0 \tag{2.13}$$

e dunque

$$\mu(\Phi^t M) = \mu(M) \tag{2.14}$$

è misura invariante per moto naturale nello spazio delle fasi.
Questo conclude la dimostrazione del teorema di Liouville. □

Corollario 2.1.2.
Se il sottoinsieme M è invariante

$$\int_M f(x)\, dpdq = \int_M f(x,t)\, dpdq \tag{2.15}$$

con f arbitraria funzione integrabile secondo Lebesgue sullo spazio delle fasi.

2.2 Il teorema di Birkhoff-Khinchin

Quello che passa sotto il nome di Teorema ergodico di Birkhoff-Khinchin (o di Birkhoff) è in realtà l'insieme di tre distinti teoremi, il primo dei quali dimostra l'esistenza quasi ovunque della media temporale di una funzione definita e sommabile in un sottoinsieme M dello spazio delle fasi lungo una traiettoria che passa per un dato punto iniziale $\overline{x}_0$; il secondo teorema assicura l'indipendenza di questo risultato dal punto inixziale $\overline{x}_0$ ed il terzo è la vera e propria asserzione della uguaglianza di tale media temporale con la media in fase, nel caso di metrica indecomponibilità (o ergodicità, in termini più moderni) come esige una adeguata fondazione della Meccanica Statistica. Il primo di questi teoremi che ora ci accingiamo ad enunciare e dimostrare consiste essenzialmente in una lunga e complessa dimostrazione per assurdo e la sua forma è dovuta a Kolmogoroff. [Si veda [1]] La dimostrazione fa anche uso del Teorema di Liouville.

Teorema 2.2.1 (Teorema di Birkhoff-Khinchin)**.**

Sia $M \subset \Gamma$ un sottoinsiema di misura $\mathcal{T}$-invariante, $\mu(\mathcal{T}^{-1}M) = \mu(M)$,ed f una funzione definita e sommabile su M. Sia $x_0 \in M$ Allora esiste q.o. in M il limite

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T f(\overline{x}_0,t)\, dt \qquad (\textit{il limite esiste anche per } T \to -\infty)$$

Si può interpretare tale limite come media temporale di $f(\overline{x}_0,t)$ lungo la traiettoria che passa per $\overline{x}_0$ durante l'intervallo di tempo $(0,T)$

Dimostrazione.
Si definiscano, per ogni intero n

$$x_n(\overline{x}_0) \stackrel{def}{=} x_n \stackrel{def}{=} \int_n^{n+1} f(\overline{x}_0, t)\, dt \tag{2.16}$$

$$y_n(\overline{x}_0) \stackrel{def}{=} y_n \stackrel{def}{=} \int_n^{n+1} \bar{f}(\overline{x}_0, t)^{-} dt \tag{2.17}$$

Vale il seguente

Lemma 2.2.2.

$\frac{1}{n} y_n(\overline{x}_0) \to 0 \quad per \quad n \to +\infty \quad q.o.\ in\ M$

Dimostrazione.
La dimostrazione è basata sul Teorema di Liouville.
Si definisca la nuova variabile di inegrazione α nell'integrale che definisce y_n: $t \stackrel{def}{=} n + \alpha$. Allora

$$y_n(\overline{x}_0) = \int_0^1 \bar{f}(\overline{x}_0, n+\alpha)^{-} d\alpha = \int_0^1 \bar{f}(\Phi^n \overline{x}_0, \alpha)^{-} d\alpha = y_0(\Phi^n \overline{x}_0) \tag{2.18}$$

perchè $f(\overline{x}_0, n+\alpha) = f(\Phi^n \overline{x}_0, \alpha)$.
Si fissi un $\epsilon > 0$ e si definiscano i sottoinsiemi $E_{n,n}, E_{n,0}$ di M

$$E_{n,n} \stackrel{def}{=} \{\overline{x}_0 \in M \mid y_n(\overline{x}_0) > \epsilon n\} = \{\overline{x}_0 \in M \mid \int_0^1 \bar{f}(\overline{x}_0, n+\alpha)^{-} d\alpha > \epsilon n\} \tag{2.19}$$

$$E_{n,0} \stackrel{def}{=} \{\overline{x}_0 \in M \mid y_0(\overline{x}_0) > \epsilon n\} = \{\overline{x}_0 \in M \mid \int_0^1 \bar{f}(\overline{x}_0, \alpha)^{-} d\alpha > \epsilon n\} \tag{2.20}$$

Durante il moto naturale in Γ, i punti di $E_{n,n}$ passano per $E_{n,0}$ durante il tempo t, perchè $\Phi^0 \overline{x}_0 \in \{\Phi^n \overline{x}_0, n \in \mathbb{Z}\}$. Infatti, nella (18), la disuguaglianza $y_n(\overline{x}_0) > n\epsilon$ è equivalente alla disuguaglianza $y_n(\Phi^n \overline{x}_0) > n\epsilon$.
Quindi $\overline{x}_0 \in E_{n,n}$ implica $\Phi^n \overline{x}_0 \in E_{n,0}$, da cui $E_{n,n} \subseteq E_{n,0}$ e viceversa $E_{n,0} \subseteq E_{n,n}$.
Il teorema di Liouville implica che $\mu(E_{n,n}) = \mu(E_{n,0})$.
Ora bisogna mostrare che la serie $\sum_{n=1}^{\infty} \mu(E_{n,n})$ converge, ovvero che converge

la serie $\sum_{n=1}^{\infty} \mu(E_{n,0})$. Sia

$$F_m \stackrel{def}{=} \{\overline{x}_0 \in M \mid m\epsilon < y_0(\overline{x}_0) \leq (m+1)\epsilon\} \tag{2.21}$$

e si osservi che è

$$E_{n,0} = \bigcup_{m=n}^{\infty} F_m \tag{2.22}$$

da cui

$$\begin{aligned}\sum_{n=1}^{\infty} \mu(E_{n,0}) &= \sum_{n=1}^{\infty}\sum_{m=n}^{\infty} \mu(F_m) = \sum_{m=1}^{\infty}\sum_{n=1}^{m} \mu(F_m) = \sum_{m=1}^{\infty} m\mu(F_m) = \frac{1}{\epsilon}\sum_{m=1}^{\infty} m\epsilon\mu(F_m) \leq \\ &\leq \frac{1}{\epsilon}\sum_{m=1}^{\infty}\int_{m\epsilon<y_0\leq(m+1)\epsilon} y_0(\overline{x}_0)d\mu \leq \frac{1}{\epsilon}\int_M y_0(\overline{x}_0)d\mu = \frac{1}{\epsilon}\int_M d\mu \int_0^1 \bar{f}(\overline{x}_0,\alpha)\bar{d}\alpha = \\ &= \frac{1}{\epsilon}\int_0^1 d\alpha \int_M \bar{f}(\overline{x}_0,\alpha)\bar{d}\mu = \frac{1}{\epsilon}\int_0^1 d\alpha\int_M \bar{f}(\overline{x}_0)\bar{d}\mu = \\ &= \frac{1}{\epsilon}\int_M \bar{f}(\overline{x}_0)d\mu < +\infty\end{aligned} \tag{2.23}$$

avendo fatto uso del teorema di Liouville e usato l'ipotesi che f è sommabile.
Segue che $\sum_{n=1}^{\infty} \mu(E_{n,0}) < +\infty$, come si voleva.
Poichè, come conseguenza del Teorema di Fubini, μ-quasi ogni $\overline{x}_0 \in M$ appartiene a non più che un numero finito della sequenza $E_{n,n} \quad n = 1, 2, ...$ si ha che per quasi ogni $\overline{x}_0 \in M \quad \exists N = N(\overline{x}_0)$ tale che per ogni $n > N(\overline{x}_0)$

$$y_n(\overline{x}_0) \leq \epsilon n \tag{2.24}$$

ovvero

$$\frac{1}{n} y_n(\overline{x}_0) \leq \epsilon \tag{2.25}$$

e per l'arbitrarietà di ϵ

$$\frac{1}{n} y_n(\overline{x}_0) \to 0 \quad per \quad n \to \infty \quad quasi\ ovunque\, in\, M \tag{2.26}$$

e ciò conclude la dimostrazione del Lemma 1. $\square$

Si definisca ora, per ogni coppia di interi $a, b,\ a < b$,

$$h_{ab}(\overline{x}_0) \overset{def}{=} \frac{1}{b-a}\int_a^b f(\overline{x}_0, t)\, dt \tag{2.27}$$

Ricordando la definizione di x_n, si ha

$$h_{ab}(\overline{x}_0) = \frac{1}{b-a}(x_a + x_{a+1} + \ldots.. + x_{b-1}) \tag{2.28}$$

$$h_{0n}(\overline{x}_0) = \frac{1}{n}(x_0 + x_1 + \ldots.. + x_{n-1}) \tag{2.29}$$

Lemma 2.2.3.
Se h_{0n}, quando $n \to \infty$ assumendo valori interi, non ha limite nell'insieme M di misura positiva $\mu(M) > 0$, allora esistono due numeri reali $\alpha, \beta,\ \alpha < \beta$, ed esiste un sottoinsieme $M' \subset M$, con $\mu(M') > 0$, tali che per ogni $\overline{x}_0 \in M'$

$$l(\overline{x}_0) \overset{def}{=} \liminf_{n\to\infty} h_{0n}(\overline{x}_0) < \alpha \tag{2.30}$$

$$L(\overline{x}_0) \overset{def}{=} \limsup_{n\to\infty} h_{0n}(\overline{x}_0) > \beta \tag{2.31}$$

Dimostrazione.
Se si considera la totalità di tutti gli intervalli (α_n, β_n) ad estremi razionali (numerati in modo qualsiasi) si ha che per $\overline{x}_0 \in M$ è $l(\overline{x}_0) < L(\overline{x}_0)$ e quindi ci sarà un intervallo (α_m, β_m) per il quale

$$l(\overline{x}_0) < \alpha_m < \beta_m < L(\overline{x}_0) \tag{2.32}$$

Denotato con M_m l'insieme di tutti i punti $\overline{x}_0 \in M$ che soddisfano tale disuguaglianza, si avrà l'unione disgiunta

$$M = \bigcup_{m=1}^{\infty} M_m \tag{2.33}$$

e per almeno un valore di m sarà $\mu(M_m) > 0$ poiche è $\mu(M) > 0$. Posto $\alpha \overset{def}{=} \alpha_m, \beta \overset{def}{=} \beta_m, M' \overset{def}{=} M_m$ il lemma 2 è dimostrato. $\square$

Si assumano ora soddisfatte le condizioni del lemma 2. Si vuole mostrare che portano ad una contraddizione (e che deve necessariamente essere

$\mu(M') = 0)$.
Sia $\overline{x}_0 \in M'$ e sia (a, b) un intervallo ad estremi interi, con $a < b$.

Definizione
Il segmento (a, b) è un segmento proprio del punto $\overline{x}_0 \in M'$ se:
1) $h_{ab}(\overline{x}_0) > \beta$
2) $h_{ab'}(\overline{x}_0) \leq \beta$ per tutti i b' tali che $a < b' < b$

Si vuole mostrare che
due segmenti propri (a_1, b_1) e (a_2, b_2) dello stesso punto $\overline{x}_0$ non possono parzialmente sovrapporsi l'uno all'altro.
Infatti, se fosse, per esempio, $a_1 < a_2 < b_1 < b_2$, si avrebbe

$$(b_1 - a_1)h_{a_1b_1} = (a_2 - a_1)h_{a_1a_2} + (b_1 - a_2)h_{a_2b_1} \tag{2.34}$$

mentre per la 2) si avrebbe

$$h_{a_1a_2} \leq \beta \quad h_{a_2b_1} \leq \beta \quad h_{a_1b_1} > \beta \tag{2.35}$$

che condurrebbe alla relazione contraddittoria

$$(b_1 - a_1)\beta < \beta(a_2 - a_1) + \beta(b_1 - a_2) = \beta(b_1 - a_1) \tag{2.36}$$

Definizione
Un segmento proprio di $\overline{x}_0$ è un segmento proprio massimale di rango s se la sua lunghezza non eccede s e se non è contenuto in nessun altro segmento proprio la cui lunghezzan non eccede s.

Allora ogni segmento proprio di lunghezza non eccedente s è contenuto in uno ed un solo segmento proprio massimale di rango s.
Infatti tra tutti i segmenti propri di lunghezza non eccedente s e contenenti il dato segmento, ve ne sarà uno di massima lunghezza. Questo sarà il segmento proprio massimale di rango s. La sua unicità deriva dal fatto che, se ce ne fossero due, allora essi, contenendo lo stesso segmento, avrebbero punti in comune. Ma allora o uno sarebbe contenuto nell'altro, e non sarebbe quindi

massimale, oppure essi dovrebbero parzialmente sovrapporsi, cosa che si è vista essere impossibile.
Per ogni $s \in \mathbb{Z}^+$ sia

$$M_s \stackrel{def}{=} \{\overline{x}_0 \in M' \mid h_{0n}(\overline{x}_0) > \beta \quad per\, almeno\, un\ n \leq s\} \tag{2.37}$$

Per s sufficientemente grande, ogni $\overline{x}_0 \in M'$ appartiene a tutti gli M_s, sicchè si può scrivere

$$M' = \bigcup_{s=1}^{\infty} M_s \tag{2.38}$$

e poichè $M_s \subset M_{s+1}$

$$\mu(M') = \lim_{s \to \infty} \mu(M_s) \tag{2.39}$$

Essendo per ipotesi, $\mu(M') > 0$, per s sufficientemente grande si avrà

$$\mu(M_{s'}) > 0 \tag{2.40}$$

per un certo intero positivo s'.

Lemma 2.2.4.
Affinchè $\overline{x}_0 \in M_{s'}$ è necessario e sufficiente che $\overline{x}_0$ abbia un segmento massimale proprio (a, b) di rango s' tale che $a \leq 0 < b$.

Dimostrazione.
$\Rightarrow$) Sia $\overline{x}_0 \in M'_s$ e sia n il più piccolo intero positivo per il quale si abbia $h_{0n}(\overline{x}_0) > \beta$ così che $n \leq s$.
Allora $(0, n)$ è un segmento proprio di $\overline{x}_0$. Come si è mostrato sopra, il segmento $(0, n)$ è contenuto in un (unico) segmento proprio massimale di rango s', (a, b) che soddisfa $a \leq 0 < b$.
$\Leftarrow$) $\overline{x}_0$ abbia un segmento proprio massimale (a, b) di rango s tale da soddisfare $a \leq 0 < b$. E' allora sufficiente mostrare che in questo caso $h_{0b}(\overline{x}_0) > \beta$ perchè $b \leq b - a \leq s'$ (essendo $a \leq 0$).
Se $a = 0$ l'asserto è ovvio: $(0, b)$ è un segmento proprio del punto $\overline{x}_0$, per il quale $h_{0b}(\overline{x}_0) > \beta$.

Se $a < 0$, si ha
$(b-a)h_{ab}(\overline{x}_0) = (x_a + ... + x_{-1}) + (x_0 + ... + x_{b-1})$, ovvero
$(b-a)h_{ab}(\overline{x}_0) = -ah_{a0}(\overline{x}_0) + bh_{0b}(\overline{x}_0)$, da cui
$h_{0b}(\overline{x}_0) = \frac{1}{b}[(b-a)h_{ab}(\overline{x}_0) + ah_{a0}(\overline{x}_0)]$-
Ma, per definizione di segmento proprio, $h_{ab}(\overline{x}_0) > \beta \quad h_{a0}(\overline{x}_0) \leq \beta$, e poichè $b - a > 0$ e $a < 0$
$h_{0b}(\overline{x}_0) > \frac{1}{b}[(b-a)\beta + a\beta] = \beta$, cioè $h_{0b}(\overline{x}_0) > \beta$.
Questo conclude la dimostrazione del Lemma 3. $\square$

Si consideri ora un qualsiasi punto $\overline{x}_0$ dell'insieme $M_{s'}$ ed un segmento proprio massimale (a, b) di rango s' corrispondente a $\overline{x}_0$ nel senso del Lemma 3, cosicchè $a \leq 0 < b, \quad b - a \leq s'$.
Si ponga $q \overset{def}{=} b - a$, e $p \overset{def}{=} -a$. Allora $1 \leq q \leq s', \quad 0 \leq p \leq q - 1$. Denotiamo $\delta_{pq} \overset{def}{=} (-p, -p + q)$
e indichiamo con M_{pq} l'insieme dei punti di $M_{s'}$ che corrispondono al segmento δ_{pq} nel senso del Lemma 3, cosicchè $M_{s'} = \bigcup_{q=1}^{s'} \bigcup_{p=0}^{q-1} M_{pq}$.
Nel moto naturale nello spazio delle fasi, l'insieme M_{0q} va a finire nell'insieme M_{pq} dopo p unità di tempo perchè $h_{0q}(\overline{x}_0) = h_{-p,q-p}(\overline{x}_{0_p})$.
Quindi $\mu(M_{pq}) = \mu(M_{0q})$, $(0 \leq p \leq q-1)$, per il Teorema di Liouville. Inoltre gli insiemi M_{pq} con differenti coppie di indici p e q non possono avere punti in comune. Infine, per ogni funzione sommabile $\phi(\overline{x}_0)$ è

$$\int_{M_{pq}} \phi(\overline{x}_0)\, dV = \int_{M_{0q}} \phi(\overline{x}_0, p)\, dV \tag{2.41}$$

Si conclude che

$$\int_{M_{s'}} dV \int_0^1 f(\overline{x}_0, t)\, dt = \sum_{q=1}^{s'} \sum_{p=0}^{q-1} \int_{M_{pq}} dV \int_0^1 f(\overline{x}_0, t)\, dt =$$
$$= \sum_{q=1}^{s'} \sum_{p=0}^{q-1} \int_{M_{0q}} dV \int_0^1 f(\overline{x}_{0_p}, t)\, dt = \sum_{q=1}^{s'} \sum_{p=0}^{q-1} \int_{M_{0q}} dV \int_p^{p+1} f(\overline{x}_0, \alpha)\, d\alpha =$$
$$= \sum_{q=1}^{s'} \int_{M_{0q}} dV \int_0^q f(\overline{x}_0, \alpha)\, d\alpha = \sum_{q=1}^{s'} \int_{M_{0q}} q h_{0q}(\overline{x}_0)\, dV >$$
$$> \beta \sum_{q=1}^{s'} q\mu(M_{0q}) = \beta \sum_{q=1}^{s'} \sum_{p=0}^{q-1} \mu(M_{pq}) = \beta\mu(M_{s'}) \tag{2.42}$$

Questa relazione vale per s' sufficientemente grande. Facendo tendere s' ad infinito, si ottiene

$$\mathcal{I} \stackrel{def}{=} \int_{M'} dV \int_0^1 f(\overline{x}_0, t)\, dt \geq \beta\mu(M') \tag{2.43}$$

Ma poichè (nel Lemma 2) per tutti i punti di M' era anche $\liminf\limits_{n\to\infty} h_{0n}(\overline{x}_0) < \alpha$ si dimostra in modo analogo che

$$\mathcal{I} \leq \alpha\mu(M') \tag{2.44}$$

Segue allora dalle relazioni contraddittorie $\begin{cases} \beta\mu(M') \leq \mathcal{I} \leq \alpha\mu(M') \\ \qquad\quad \alpha < \beta \end{cases}$
che deve necessariamente essere $\mu(M') = 0$ (cioè l'assunzione $\mu(M') > 0$ non è possibile). In altre parole (dal Lemma 2 segue che) il limite $\lim_{n\to\infty} h_{0n}(\overline{x}_0)$ deve esistere quasi ovunque.
Per completare la dimostrazione del teorema, occorre rimuovere l'assunzione che il parametro n assuma solo valori interi. Ciò segue dal Lemma 1. Infatti, indicata con $[b]$ la parte intera di $b \in \mathbb{R}$, poichè l'espressione $\dfrac{1}{b}\displaystyle\int_0^{[b]} f(\overline{x}_0, t)\, dt$, per $b \to \infty$, differisce da $\dfrac{1}{[b]}\displaystyle\int_0^{[b]} f(\overline{x}_0, t)\, dt$ solo per un fattore che tende a zero (perchè $\mid \frac{1}{b} - \frac{1}{[b]} \mid = \mid \frac{[b]-b}{b[b]} \mid \leq \frac{1}{b[b]} \to 0$ per $b \to \infty$), e poichè questa ha un limite quasi ovunque, anche il limite $\lim\limits_{b\to\infty} \dfrac{1}{b}\displaystyle\int_0^{[b]} f(\overline{x}_0, t)$ esiste quasi ovunque.
D'altra parte

$$\begin{aligned} \mid \frac{1}{b}\int_0^b f(\overline{x}_0, t)\, dt - \frac{1}{b}\int_0^{[b]} f(\overline{x}_0, t)\, dt \mid &= \mid \frac{1}{b}\int_{[b]}^b f(\overline{x}_0, t)\, dt \mid \leq \\ \leq \frac{1}{b}\int_{[b]}^b \mid f(\overline{x}_0, t) \mid\, dt &\leq \frac{1}{[b]}\int_{[b]}^{[b]+1} \mid f(\overline{x}_0, t) \mid\, dt = \\ &= \frac{y_{[b]}(\overline{x}_0)}{[b]} \to 0, \qquad b \to \infty \end{aligned} \tag{2.45}$$

in virtù del Lemma 1. Da cui il limite

$$\lim_{b\to\infty} \frac{1}{b}\int_0^b f(\overline{x}_0, t)\, dt \tag{2.46}$$

pure esiste quasi ovunque.
Questo completa la dimostrazione del Teorema di Birkhoff-Khinchin. □

2.3 Indipendenza del limite $\overline{f}(\overline{x}_0)$ dal punto $\overline{x}_0$

Il limite $\overline{f}(\overline{x}_0)$ del teorema di Birkhoff-Khinchin rappresenta la media temporale della funzione $f(\overline{x}_0)$ lungo la traiettoria $x = \Phi^t \overline{x}_o$ che passa per il punto $\overline{x}_0$, che d'ora in poi indicheremo con x_0.
Ma occorre ancora dimostrare che tale media non dipende dal punto x sulla traiettoria che passa per x_0.

Teorema 2.3.1.

$$\textit{Per ogni} \quad x = \Phi^t x_0 \qquad \overline{f}(x) = f(x_0)$$

Dimostrazione.
Si prenda $t > 0$. Per ipotesi esiste il limite $\lim_{T\to\infty} \frac{1}{T+t} \int_0^{T+t} \int_0^{T+t} f(x_0, \alpha \, d\alpha$.
Poichè la differenza

$$\frac{1}{T}\int_0^{T+t} f(x_0,\alpha)\,d\alpha - \frac{1}{T+t}\int_0^{T+t} f(x_0.\alpha)d\alpha = \frac{t}{T(T+t)}\int_0^{T+t} f(x_0,\alpha)\,d\alpha \to 0, \;\; t \to \infty \tag{2.47}$$

si ha pure

$$\overline{f}(x_0) = \lim_{T\to\infty} \frac{1}{T}\int_0^{T+t} f(x_0,\alpha)\,d\alpha \tag{2.48}$$

Ma allora

$$\begin{aligned} \overline{f}(\Phi^t x_0) = \overline{f}(x) = \frac{1}{T}\int_0^T f(x,\alpha)\,d\alpha = \frac{1}{T}\int_0^T f(x_0, t+\alpha)\,d\alpha = \\ = \frac{1}{T}\int_t^{T+t} f(x_0,\alpha')\,d\alpha' = \frac{1}{T}\int_0^{T+t} f(x_0,\alpha')\,d\alpha' - \frac{1}{T}\int_0^t f(x_0,\alpha')\,d\alpha' \to \\ \to \overline{f}(x_0) - 0 = \overline{f}(x_0), \qquad T \to \infty \end{aligned} \tag{2.49}$$

Dunque

$$\overline{f}(x) = \overline{f}(x_0) \qquad \textit{per ogni} \;\; x = \Phi^t x_0 \quad \textit{e per ogni}\; t. \tag{2.50}$$

Il teorema è dimostrato. □

2.4 Indecomponibilità metrica e uguaglianza tra media temporale e media in fase

Il caso più importante del Teorema di Birkhoff-Khinchin si ha quando il sottoinsieme dello spazio delle fasi $M \in \Gamma$ invariante, $\mu(\Phi^t M) = \mu(M)$, è metricamente indecomponibile, cioè non può essre spezzato in
$M = M_1 \cup M_2 \quad \text{con} \quad \mu(M_i) = \mu(\Phi^t M_i) \quad \text{e} \quad \mu(M_i) > 0 \quad i = 1, 2.$
Questo significa quanto segue: M, in quanto insieme invariante, è un certo insieme completo di traiettorie. Se questo insieme completo di traiettorie fosse separabile in due insiemi completi di traiettorie M_1 e M_2, solo i seguenti due casi sarebbero possibili:
1) M_1 e M_2 non sono misurabili
2) uno dei due, diciamo M_2, ha misura nulla $\mu(M_2) = 0$ e $\mu(M_1) \equiv \mu(M)$-
Nel caso in cui M èmetricamente indecomponibile, il teorema di Birkhoff-Khinchin può essere reso più preciso. Precisamente vale il seguente

Teorema 2.4.1.
Se M è metricamente indecomponibile, allora si ha:

$$\overline{f}(x) = \frac{1}{\mu(M)} \int_M f(x)\, d\mu \stackrel{def}{=} < f >$$

Questo teorema asserisce che, nel caso di metrica indecomponibilità dell'insieme $M \subset \Gamma$, la media temporale $\overline{f}(x)$ di una qualsiasi funzione sommabile f
- è la stessa per quasi ogni punto iniziale x_0
-coincide con la media in fase $< f >$ della stessa funzione f.

Dimostrazione.
La dimostrazione consiste di due parti.
(Prima parte) Si dimostra che $\quad \overline{f}(x) = cost. \qquad \mu - q.o.$
Se così non fosse, esisterebbe un numero reale α che, nel separare M in due parti M_1 e M_2, definiti rispettivamente dalle condizioni

$$\overline{f})(x) > \alpha \quad su\, M_1, \quad \overline{f}(x) \leq \alpha \quad su\, M_2$$

avremmo $\mu(M_1) > 0$ e $\mu(M_2) > 0$ con $M = M_1 \cup M_2$.
Per provare ciò, per ogni $n \in \mathbb{Z}$, si suddivida $\mathbb{R}$ in intervalli della forma

$(\frac{k}{2^n}, \frac{k+1}{2^n})$, $k \in \mathbb{Z}$, che sono tutti disgiunti. Chiamiamo essnziale uno di tali intervalli δ_{nk} se

$$\mu\big\{x_0 \in M \mid \overline{f}(x) \in \delta_{nk}\big\} > 0$$

Se esistessero due di tali intervalli, l'esistenza dell'α di cui sopra sarebbe provata, ma avremmo una contraddizione con l'ipotesi di metrica indecomponibilità.
Se tuttavia esiste un solo δ_n essenziale, per ogni n, allora è $\delta_{n+1} \subset \delta_n$ e così la successione di intervalli $\big\{\delta_n\big\}_{n=1}^{\infty}$ ha un solo punto comune, che è $\lim_{n\to\infty} \delta_n = \alpha$ $(k \in \mathbb{Z})$. In questo caso è ovvio che si può affermare che

$$\overline{f}(x) = \alpha \overset{def}{=} a = cost. \qquad quasi\ ovunque\, in\, M.$$

(Seconda parte) Resta da dimostrare che

$$a = \frac{1}{\mu(M)} \int_M f(x) d\mu =< f >$$

A tale scopo sia

$$f_T(x) \overset{def}{=} \frac{1}{T} \int_0^T f(x,t)\, dt \tag{2.51}$$

Allora si ha identicamente

$$a = \frac{1}{\mu(M)} \int_M a d\mu = \frac{1}{\mu(M)} \int_M [a - f_T(x)] d\mu + \frac{1}{\mu(M)} \int_M f_T(x) d\mu \tag{2.52}$$

Dall'invarianza di M segue

$$\frac{1}{\mu(M)} \int_M f_T(x)\, d\mu = \frac{1}{T\mu(M)} \int_0^T dt \int_M f(x,t)\, d\mu = \frac{1}{T\mu(M)} \int_0^T dt \int_M f(x)\, d\mu = \\ = \frac{1}{\mu(M)} \int_M f(x)\, d\mu \overset{def}{=} < f >$$

Da cui

$$a = \frac{1}{\mu(M)} \int_M [a - f_T(x)]\, d\mu + < f > \tag{2.53}$$

e la quantità

$$\frac{1}{\mu(M)} \int_M [a - f_T(x)]\, d\mu = a - < f > \tag{2.54}$$

non dipende da t.
Il teorema sarà dimostrato se si prova che questo integrale è nullo.
A tale scopo, sia $\epsilon > 0$ arbitrariamente piccolo. Siano

$$M_1(T) \stackrel{def}{=} \left\{x \in M \mid \mid a - f_T(x) \mid < \epsilon\right\} \quad e \quad M_2(T) \stackrel{def}{=} M \setminus M_1(T)$$

Si ha

$$\begin{aligned}\mid \int_M [a - f_T(x)]\, d\mu \mid \leq \int_{M_!(T)} \mid a - f_T(x) \mid d\mu + \int_{M_2(T)} \mid [a - f_T(x)] d\mu \leq \\ \leq \epsilon\mu(M_1) + \mid a \mid \mu(M_2(T)) + \int_{M_2(T)} \mid f_T(x) \mid d\mu\end{aligned} \tag{2.55}$$

Ora, poichè $f_T(x) \to a$, per $T \to \infty$, q.o. in M, $\mu(M_2(T)) \to 0$ per $T \to \infty$, prt T sufficientemente grande si ha $\mu(M_2(T)) < \epsilon$. Ma

$$\int_{M_2(T)} \mid f_T(x) \mid d\mu \leq \frac{1}{T} \int_0^T dt \int_{M_2(T)} \mid f(x,t) \mid d\mu = \frac{1}{T} \int_0^T dt \int_{M_2(T,t)} \mid f(x) \mid d\mu \tag{2.56}$$

dove $M_2(T,t)$ è l'insieme in cui $M_2(T)$ va a finire durante l'intervallo di tempo t, nel moto naturale nello spazio delle fasi. Quindi dal teorema di Liouville

$$\mu(M_2(T,t) = \mu(M_2(T))$$

e $\mu(M_(T,t)) \to 0$ per $T \to \infty$, uniformemente rispetto a t. In virtù dell'assoluta continuità degli integrali di funzioni sommabili, si può prendere T tanto grande che, per tutti i t

$$\int_{M_2(T,t)} \mid f(x) \mid d\mu < \epsilon$$

La (57) mostra che, in questo caso,

$$\int_{M_2(T)} \mid f_T(x) \mid d\mu < \epsilon$$

e allora la (56) dà

$$\mid \int_M [a - f_T(x)] d\mu \mid \leq \epsilon\mu(M_1) + \mid a \mid \epsilon + \epsilon$$

in cui il primo membro sarà arbitrariamente picoolo per T sufficientemente grande e, poichè non dipende da T, deve essere uguale a zero:

$$\mid \int_M [a - f_T(x)] d\mu \mid = 0 \tag{2.57}$$

come era da dimostrarsi. □

Questo conclude la dimostrazione di

$$\overline{f}(x) = cost. =< f > \quad q.o. \quad ovvero \quad \lim_{T\to\infty}\int_0^T f(x)dt = \frac{1}{\mu(M)}\int_M f(x)d\mu = cost. \in \mathbb{R}.$$

Capitolo 3

Il teorema ergodico medio di von Neumann

3.1 Il Teorema di ricorrenza di Poincarè

Questa proprietà generale della dinamica discende solo dall'esistenza di una distribuzione di probabilità T-invariante:

$$\mu(T^{-1}A) = \mu(A) \tag{3.1}$$

Teorema 3.1.1 (Teorema di ricorrenza di Poincarè)**.**
Per ogni insieme A, di misura positiva $\mu(A) > 0$, l'orbita di ogni punto $x \in A$ ritorna con certezza infinite volte nell'insieme A:

$$(\Phi^t x) \in A \quad \forall t > t_0 \quad con \quad \Phi^{t_0} x = x \tag{3.2}$$

Cioè

$$\mu(\{x \in A | \Phi^t x \in A \quad per\, un\, numero\, finito\, di\, t\}) = 0 \tag{3.3}$$

In altro modo, scelto a caso un x_0

$$\mathbb{P}\{\Phi^t x_0 \in A, \quad per\ infiniti\ t\} = 1 \tag{3.4}$$

Questo significa che, per una generica funzione f, la funzione $f(\Phi^t x_0)$ non converge a nulla, in quanto i suoi valori continuano ad oscillare incessantemente, poichè tornano ad avere ogni volta il valore che avevano inizialmente.

Diamo la dimostrazione per mappe discrete, che risulta più semplice e chiara.

Dimostrazione. Sia

$$N_1 \stackrel{def}{=} \{x \in A | \Phi^n(x) \notin A \quad \forall n\} \tag{3.5}$$

l'insieme dei punti che non ritornano in A. Si vuole mostrare che $\mu(N_1) = 0$.
Gli insiemi $\Phi^n N_1$, evoluti di N_1 al variare di $n \in \mathbb{N}$, sono due a due disgiunti. Infatti, se esistesse $x \in \Phi^n N_1 \cap \Phi^m N_1$ allora seguirebbe che $\Phi^{-m}x \in \Phi^{n-m}N_1 \cap N_1$, cioè $\Phi^{-m}x$ sarebbe un punto di N_1 che ritorna, dopo $n - m$ iterazioni, di nuovo in N_1 cioè in A perchè N_1 è, per costruzione, un sottoinsieme di A, contro l'ipotesi che i punti di N_1 non ritornano mai in A.
Poichè Φ conserva la misura, è $\mu(\Phi^n N_1) = \mu(N_1)$.
Si vuole ora mostrare che $\mu(N_1) = 0$.
Consideriamo l'unione $\bigcup_n \Phi^n N_1$ degli evoluti $\Phi^n N_1$ di N_1. Poichè tali insiemi sono due a due disgiunti, la misura di $\bigcup_n \Phi^n N_1$ (che è finita, minore di $\mu(M) = 1$) deve essere uguale alla somma delle misure dei singoli evoluti. E si ha quindi:

$$1 \geq \mu\left(\bigcup_{n=0}^{\infty} \Phi^n N_1\right) = \sum_{n=0}^{\infty} \mu\left(\Phi^n N_1\right) = \sum_{n=0}^{\infty} \mu(N_1) \tag{3.6}$$

Segue che deve necessariamente essere $\mu(N_1) = 0$.
Definendo ora $A_1 \stackrel{def}{=} A \setminus N_1$, ovvero l'insieme dei punti che ritorna almeno una volta, risulta dimostrato che

$$\mu(A_1) = \mu(A) \setminus \mu(N_1) = \mu(A) \tag{3.7}$$

Iterando questo procedimento, definiamo, per ogni intero k, l'insieme N_k dei punti di A_{k-1} che non ritornano in A_{k-1} e definiamo l'insieme $A_k \stackrel{def}{=} A_{k-1} \setminus N_k$ dei punti che ritornano almeno k volte in A.
Con ragionamento analogo al precedente, si trova che $\mu(N_k) = 0$, e dunque, $\mu(A_k) = \mu(A)$.
Ma allora l'insieme $A_\infty \stackrel{def}{=} A \setminus \bigcup_k N_k$ risulta essere l'insieme dei punti che ritornano un numero infinito di volte in A. Poichè l'unione numerabile di insiemi dimisura nulla $\mu\left(\bigcup_k N_k\right) = 0$, segue che

$$\mu(A_\infty) = \mu(A) \tag{3.8}$$

cioè la tesi. $\square$

3.2 Il teorema di Kac

Sia Φ un diffeomorfismo dello spazio delle fasi $\mathcal{M}$ con misura invariante μ, che sia ergodico.
Fissato un sottoinsieme $A \subset \mathcal{M}$, per ogni $x \in A$ si indichi con $n(x)$ il tempo di primo ritorno di $\Phi^n x$ in A per l'orbita con dato iniziale x, cioè il più piccolo intero n per cui si abbia $(\Phi^n x) \in A$.
Allora si ha (Teorema di Kac, 1947):

$$\int_A n(x)\, d\mu = 1 \tag{3.9}$$

Questa relazione si può leggere, dividendo ambo i membri per $\mu(A)$, come

$$< n(x) >= \frac{\int_A n(x)\, d\mu}{\int_A d\mu} = \frac{1}{\int_A d\mu} \tag{3.10}$$

e si può rileggere dicendo che
il tempo medio di primo ritorno in A è inversamente proporzionale alla misura dell'insieme A.
Più l'insieme A è piccolo, più i tempi si allungano. E viceversa.

3.3 L'evoluzione come operatore unitario sulle osservabili: il teorema di Koopman

Per dimostrare il Teorema ergodico medio di von Neumann occorre cambiare prospettiva:
invece di concentrarsi sul flusso nello spazio delle fasi (cheè inosservabile per ipotesi), ci si focalizza sulle evoluzioni delle osservabili, cioè sul cambiamento delle variabili dinamiche dovuto alla evoluzione del sistema.
Qui entra in gioco lo spazio di Hilbert perchè ci si concentra sullo spazio delle funzioni a modulo quadro sommabile:

$$\int_M |\, f(x)\,|^2\, d\mu < +\infty \tag{3.11}$$

Le funzioni a modulo quadro sommabile sono significative anche perchè, per tali funzioni, la media temporale $\overline{f}(x_0)$ risulta avere finiti non solo la media, ma anche lo scarto quadratico medio.
Infatti, poichè M ha misura finita, $\int_M d\mu < +\infty$, risulta che le funzioni a modulo quadrato sommabile sono anche sommabili, in virtù della disuguaglianza di Schwarz che (per $g(x) \equiv 1(x)$) porge:

$$\int_M |\, f(x) \,|\ d\mu \leq \left[\int_M |\, f(x) \,|^2\ d\mu \right]^{1/2} \left[\int_M d\mu \right]^{1/2} \leq \|f\|_2 < +\infty \quad (3.12)$$

Dunque esistono finite la media

$$< f > \stackrel{def}{=} \int_M f(x)\, d\mu \quad (3.13)$$

ed anche lo scarto quadratico medio

$$\Delta_f^2 \stackrel{def}{=} \int_M f^2(x)\, d\mu - \left[\int_M f(x)\, d\mu \right]^2 \quad (3.14)$$

Nel seguito considereremo solo flussi discreti, cioè orbite generate da una mappa iterata.

Definizione 3.1 (Operatore di evoluzione $\hat{\mathcal{U}}_n$).
L'operatore $\hat{\mathcal{U}}_n$ di evoluzione delle funzioni è definito da

$$\hat{\mathcal{U}}_n f(x) \stackrel{def}{=} f(\Phi^n x) \quad (3.15)$$

Questo operatore gode della proprietà gruppale (additiva) perchè

$$\hat{\mathcal{U}}_{n+m} f(x) \stackrel{def}{=} f(\Phi^{n+m} x) = f(\Phi^n(\Phi^m x)) = \hat{\mathcal{U}}_n(f(\Phi^m x)) = \hat{\mathcal{U}}_n(\hat{\mathcal{U}}_m f(x)) = (\hat{\mathcal{U}}_n \hat{\mathcal{U}}_m) f(x) \quad (3.16)$$

$$\Longrightarrow \hat{\mathcal{U}}_{n+m} = \hat{\mathcal{U}}_n \hat{\mathcal{U}}_m \quad (3.17)$$

Questo implica che $\hat{\mathcal{U}}_n = \hat{\mathcal{U}}_1^n$, cioè l'evoluzione al tempo n si ottiene iterando n volte la trasformazione $\hat{\mathcal{U}} \stackrel{def}{=} \hat{\mathcal{U}}_1$

Altre proprietà

1. $\hat{\mathcal{U}}$ è lineare:
data $h(x) = \alpha f(x) + \beta g(x)$ si ha

$$\hat{\mathcal{U}}h(x) = h(\Phi x) = \alpha f(\Phi x) + \beta g(\Phi x) = \alpha\hat{\mathcal{U}}f(x) + \beta\hat{\mathcal{U}}g(x) \tag{3.18}$$

Conseguenza: mentre l'operatore di evoluzione del sistema nello spazio delle fasi è non lineare, l'evoluzione nello spazio delle osservabili è lineare.
L'operatore $\hat{\mathcal{U}}$ è inoltre invertibile perchè lo è Φ, cioè si ha

$$\hat{\mathcal{U}}^{-1}f(x) = f(\Phi^{-1}x) \tag{3.19}$$

2. L'operatore $\hat{\mathcal{U}}$ è unitario (isometrico), cioè conserva le norme delle funzioni.
Ciò segue dal fatto che μ è invariante per Φ:

$$\|\hat{\mathcal{U}}f\|^2 = \int_M |\,\hat{\mathcal{U}}f(x)\,|^2\; d\mu = \int_M |\,f(\Phi x)\,|^2\; d\mu = \int_M |\,f(y)\,|^2\; d\mu = \|f\|^2 \tag{3.20}$$

avendo operato il cambiamento di variabile $y \overset{def}{=} \Phi x$ e avendo usato l'invarianza di μ.
Ciò mostra che $\hat{\mathcal{U}}$ è isometrico ed essendo invertibile è anche unitario (in spazi finito-dimensionali)
Quanto detto dimostra il seguente

Teorema 3.3.1 (Teorema di Koopman (1930)).
Nello spazio delle osservabili l'evoluzione è data da un gruppo ad un parametro di operatori unitari

3.4 Il teorema ergodico medio di von Neumann

Usando l'operatore unitario $\hat{\mathcal{U}}$, la media temporale $\overline{f}$ di una funzione f si scrive

$$\overline{f}(x_0) = \frac{1}{n}\sum_{k=0}^{n-1} f(\Phi^k x_0) \overset{def}{=} \frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k f(x_0) = \left(\frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k\right) f(x_0) \tag{3.21}$$

Quindi lo studio della media temporale $\overline{f}(x_0)$ è ricondotto allo studio di $\frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k$, cioè allo studio della media geometrica dell'operatore $\hat{\mathcal{U}}$.
Si ha il seguente

Teorema 3.4.1 (Teorema ergodico medio di von Neumann).
L'operatore $\hat{\mathcal{S}}_n \overset{def}{=} \frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k$ *converge, in* L^2*, per* $n \to +\infty$*, all'operatore* $\hat{\mathbb{P}}_H$ *di proiezione ortogonale sul sottospazio* H *delle funzioni* f *invarianti per* Φ*:*

$$\hat{\mathcal{S}}_n \overset{def}{=} \frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k \longrightarrow \hat{\mathbb{P}}_H \quad \textit{per } n \to +\infty \tag{3.22}$$

dove H *è il sottospazio delle funzioni* Φ*-invarianti:*

$$H \overset{def}{=} \big\{f \mid \hat{\mathcal{U}}^n f(x) = f(x)\big\} = \big\{f \mid f(\Phi^n x) = f(x)\big\} \tag{3.23}$$

Dimostrazione.
Nel caso in cui $f(x)$ sia invariante per $\hat{\mathcal{U}}$, cioè $\hat{\mathcal{U}}f(x) = f(x)$, si ha

$$\hat{\mathcal{S}}_n f(x) = \frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k f(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(x) = f(x) \tag{3.24}$$

da cui

$$\hat{\mathcal{S}}_n f(x) \longrightarrow f(x) \quad \textit{ovvero} \quad \hat{\mathcal{S}}_n \longrightarrow \hat{\mathbb{P}}_H, \quad \textit{per } n \to +\infty \tag{3.25}$$

Nel caso in cui $f(x)$ può essere scritta nella forma

$$f(x) = \big(\hat{1} - \hat{\mathcal{U}}\big)g(x) \quad \textit{per una certa} \quad g(x) \tag{3.26}$$

allora, osservando che

$$\left(\frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k\right)\left(\hat{1} - \hat{\mathcal{U}}\right) = \frac{1}{n}\sum_{k=0}^{n}\left(\hat{\mathcal{U}}^k - \hat{\mathcal{U}}^{k+1}\right) = \frac{1}{n}\left(\hat{1} - \hat{\mathcal{U}}^n\right) \tag{3.27}$$

si ha

$$\frac{1}{n}\sum_{k=0}^{n-1}\hat{\mathcal{U}}^k f(x) = \frac{1}{n}\left(\hat{1} - \hat{\mathcal{U}}^n\right)g(x) \tag{3.28}$$

che tende a zero in norma di L^2, per $n \to +\infty$, perchè

$$\frac{1}{n}\|(\hat{1} - \hat{\mathcal{U}}^n)g(x)\| \leq \frac{1}{n}\left(\|g(x)\| + \|\hat{\mathcal{U}}^n(g(x))\|\right) \leq \frac{1}{n}2\|g(x)\| \tag{3.29}$$

essendo $\hat{\mathcal{U}}$ isometrico: $\|\hat{\mathcal{U}}(g)\| = \|g\|$.
Infine, il sottospazio H delle funzioni invarianti è il complemento ortogonale $H'^{\perp}$ del sottospazio H' di tutte le funzioni $f(x)$ che si possono scrivere nella forma $f(x) = (\hat{1} - \hat{\mathcal{U}})g(x)$

$$\mathcal{H}' \stackrel{def}{=} \left\{f(x) \in L^2 \mid f(x) = (\hat{1} - \hat{\mathcal{U}})g(x), g \in L^2\right\} \tag{3.30}$$

$$\mathcal{H} = \mathcal{H}'^{\perp} \tag{3.31}$$

Infatti
$\mathcal{H}'^{\perp}$ è costituito da tutte le funzioni $h(x)$ per le quali $< h, f >= 0 \quad \forall f \in \mathcal{H}'$ ovvero $< h, (\hat{1} - \hat{\mathcal{U}})g >= 0$ cioè, poichè $\hat{\mathcal{U}}$ è unitario, $\hat{\mathcal{U}}^{-1} = \hat{\mathcal{U}}^{\dagger}$, $< (\hat{1} - \hat{\mathcal{U}}^{-1})h, g >= 0$, per ogni g, e poichè il prodotto scalare è non degenere, $(\hat{1} - \hat{\mathcal{U}}^{-1})h(x) = 0$, ovvero $h(x) = \hat{\mathcal{U}}h(x)$ ($h(x)$ è invariante), ovvero $h(x) \in \mathcal{H}$.
Ma in uno spazio di Hilbert ogni f si può decomporre in

$$f(x) = h(x) + g(x) \qquad h \in \mathcal{H},\ g \in \mathcal{H}^{\perp} \tag{3.32}$$

Siccome si è visto che per le funzioni $h(x)$ il limite è $h(x)$ stessa, mentre per le funzioni $g(x)$ il limite è nullo, si ha che

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \hat{\mathcal{U}}^k f(x) = h(x) \qquad h(x) \in \mathcal{H} \tag{3.33}$$

ovvero

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \hat{\mathcal{U}}^k = \hat{\mathbb{P}}_H \tag{3.34}$$

E questo conclude la dimostrazione del teorema ergodico medio di von Neumann. $\square$

Capitolo 4

Il Teorema ergodico quantistico di von Neumann

4.1 Preliminari alla dimostrazione del teorema ergodico quantistico

4.1.1 La formulazione quantomeccanica della Meccanica Statistica di Giibbs

Von Neumann parte dal presopposto che tutte le misurazioni macroscopiche sul sistema in esame che sono possibili, siano possibili simultaneamente. Dunque devono esistere degli operatori $\hat{A}, \hat{B}, \hat{C}$,...., che commutano tra loro: $[\hat{A}, \hat{B}] = [\hat{A}, \hat{C}] = [\hat{B}, \hat{C}] = \cdots = 0$, e che possiedono un comune sistema ortonormale completo di autofunzioni ω_1, ω_2,.... .
Ma ci si deve aspettare che, tra tali autofunzioni, ci siano gruppi di molte ω_n su cui ogni operatore macroscopico $\hat{A}, \hat{B}, \hat{C}, ...$,abbia lo stesso autovalore, perchè, se così non fosse, una simultanea osservazione delle grandezze macroscopiche $\mathcal{A}, \mathcal{B}, \mathcal{C}, ...$, permetterebbe di distinguere con esattezza tra i microstati $\omega_1, \omega_2, ...$, cosa che è impossibile.
Quindi esistono gruppi di autofunzioni degeneri per tutte le osservabili macroscopiche.
Tali gruppi vengono indicati da von Neumann con due indici, $\omega_{\lambda,p}$, in modo tale da avere

$$\{\omega_{1,p}, \omega_{2,p,...,\omega_{s_p,p}}\} \quad per \quad p = 1, 2, ... \quad e \quad \lambda = 1, 2, ..., s_p \tag{4.1}$$

Cioè, se w_p, $(p = 1, 2, ...)$, è un autovalore degenere per $\hat{A}, \hat{B}, \hat{C}, ...$, allora s_p è la sua molteplicità di degenerazione (molteplicità geometrica dell'autovalore).
Se tutti gli stati del p-esimo gruppo degenere vengono mescolati con pesi uguali (pari a s_p), si ottiene un ensemble statistico con operatore statistico $\hat{\rho}_p$ definito come

$$\hat{\rho}_p \stackrel{def}{=} \frac{1}{s_p}\hat{\mathbb{E}} \stackrel{def}{=} \frac{1}{s_p}\sum_{\lambda=1}^{s_p} \hat{\mathbb{P}}_{\omega_{\lambda,p}} = \frac{1}{s_p}\sum_{\lambda=1}^{s_p} |\ \omega_{\lambda,p} >< \omega_{\lambda,p}\ | \tag{4.2}$$

Infatti si ha $Tr(\hat{\rho}_p) = 1$, poichè

$$Tr(\hat{\rho}_p) = \frac{1}{s_p}Tr(\hat{\mathbb{E}}_p) = \frac{1}{s_p}\sum_{\lambda=1}^{s_p} Tr(\hat{\mathbb{P}}_{\omega_{\lambda,p}}) = \frac{1}{s_p}\sum_{\lambda=1}^{s_p} Tr(|\ \omega_{\lambda,p} >< \omega_{\lambda,p}\ |) = \frac{1}{s_p}s_p\, 1 = 1 \tag{4.3}$$

Inoltre, ogni operatore macroscopico, ad es. $\hat{A}$, è tale che

$$\hat{A} \quad \omega_{\lambda,p} = w_p \quad \omega_{\lambda,p} \quad per \quad \lambda = 1, ..., s_p \quad e \quad p = 1, 2, ... \tag{4.4}$$

e

$$\hat{A} = \sum_p w_p \sum_{\lambda=1}^{s_p} |\ \omega_{\lambda,p} >< \omega_{\lambda,p}\ | = \sum_p w_p \sum_{\lambda=1}^{s_p} \hat{\mathbb{P}}_{\omega_{\lambda,p}} = \sum_p w_p \hat{\mathbb{E}}_p \tag{4.5}$$

cioè è combinazione lineare degli $\hat{\mathbb{E}}_p$, con gli autovalori w_p come coefficienti.

- Il significato di $\frac{1}{s_p}\hat{\mathbb{E}}_p$

$\frac{1}{s_p}\hat{\mathbb{E}}_p$ può essere visto come l'operatore statistico dell'ensemble in cui tutte le grandezze macroscopiche $\mathcal{A}, \mathcal{B}, \mathcal{C}$,... hanno valori corrispondenti al p-esimo

gruppo (per p=1,2,...), dove s_p è il numero di microstati aventi lo stesso peso.
$\frac{1}{s_p}\hat{\mathbb{E}}_p$ è dunque la p-esima alternativa, tra quelle concernenti le proprietà del sistema, che può essere distinta attraverso una misura macroscopica.
Quindi $\frac{1}{s_p}\hat{\mathbb{E}}_p$ è, per von Neumann, l'equivalente della cella di fase della meccanica statistica di Gibbs, e la sua dimensione $s_p = Tr(\hat{\mathbb{E}}_p)$ è il numero di effettivi microstati in tale cella, dunque è una misura della grossolanità della prospettiva macroscopica.

Il problema della grandezza microscopica $\hat{H}$ e la sua riduzione a grandezza macroscopica

Sia $\hat{H}$ l'operatore energia,dotato di autofunzioni $\phi_1, \phi_2, ...$, ed autovalori $W_1, W_2, ...$, così che (in decomposizione spettrale)

$$\hat{H} = \sum_n W_n \hat{\mathbb{P}}_{\phi_n} = \sum_n W_n \mid \phi_n >< \phi_n \mid \tag{4.6}$$

Questa non è una approssimazione macroscopica, poichè $\hat{H}$ è l'esatta energia microscopica del sistema.
Il problema è che i $\mid \phi_n >$ sono differenti dagli $\mid \omega_{\lambda,p} >$ e $\hat{H}$ non può essere una combinazione lineare degli $\hat{\mathbb{E}}_p$.
Von Neumann qui ricorre ad un espediente che consiste nel considerare una ridotta accuratezza, raggruppando gli autovalori $W_1, W_2, ...$ in gruppi

$$\{W_{1,a}, W_{2,a}, ..., W_{S_a,a}\} \quad a = 1, 2, ... \tag{4.7}$$

scrivendo, al posto di W_n e ϕ_n,

$$W_{\rho,a} \quad e \quad \phi_{\rho,a}, \quad con \quad a = 1, 2, ... \quad e \quad \rho = 1, 2, ..., S_a \tag{4.8}$$

in modo tale che tutti i $W_{\rho,a}$ con lo stesso a siano macroscopicamente vicini l'uno all'altro e solo i gruppi con differenti a possano essere macroscopicamente distinguibili l'uno dall'altro.
Per poter giustificare il fatto che l'intero gruppo $\{W_{1,a}, ..., W_{S_a,a}\}$ (con a fissato) sia macroscopicamente misurabile, von Neumann definisce la funzione seguente

$$f_a(x) = \begin{cases} 1, & per \quad x = W_{1,a}, W_{2,a}, ..., W_{S_a,a} \\ 0, & altrimenti \end{cases} \qquad (a\, fissato) \tag{4.9}$$

La funzione operatoriale $f_a(\hat{H})$ assume il valore 1 quandoi il valore dell'energia appartiene al suddetto gruppo ed è zero altrimenti.
Quindi è una quantità misurabile macroscopicamente.
Da $\hat{H} = \sum_n W_n \hat{\mathbb{P}}_{\phi_n}$ segue

$$f_a(\hat{H}) = \sum_n f_a(W_n)\hat{\mathbb{P}}_{\phi_n} = \sum_a \sum_{\rho=1}^{S_a} f_a(W_{\rho,a})\hat{\mathbb{P}}_{\phi_{\rho,a}} = \sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}} \tag{4.10}$$

e questa quantità deve ora essere combinazione lineare degli $\hat{\mathbb{E}}_p$.
Ora, poichè gli operatori

$$\hat{\mathbb{P}} \stackrel{def}{=} \sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}} \quad e \quad \hat{\mathbb{E}}_p = \sum_{\lambda=1}^{s_p} \hat{\mathbb{P}}_{\omega_{\lambda,p}} \tag{4.11}$$

soddisfano $\hat{\mathbb{P}}^2 = \hat{\mathbb{P}}$ e $\hat{\mathbb{E}}_p^2 = \hat{\mathbb{E}}_p$, e inoltre $\hat{\mathbb{E}}_{p'}\hat{\mathbb{E}}_p = 0$ per $p' \neq p$, i coefficienti c di tali combinazioni lineari devono soddisfare l'equazione $c^2 = c$, cioè devono essere uguali a 0 oppure a 1.
Ciò significa che $\sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}}$ è la somma di solo alcuni degli $\hat{\mathbb{E}}_p$, che von Neumann chiama

$$\hat{\mathbb{E}}_{1,a}, \hat{\mathbb{E}}_{2,a}, ..., \hat{\mathbb{E}}_{N_a,a} \quad (\textit{abbreviato} \quad \textit{con} \quad \hat{\mathbb{E}}_{\nu,a}, \quad \textit{per} \quad \nu = 1, ..., N_a) \tag{4.12}$$

Così

$$\sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}} = \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a}. \tag{4.13}$$

Considerando la traccia di ambo i membri, si ha

$$\sum_{\rho=1}^{S_a} Tr(\hat{\mathbb{P}}_{\phi_{\rho,a}}) = S_a 1 = \sum_{\nu=1}^{N_a} Tr(\hat{\mathbb{E}}_{\nu,a}) = \sum_{\nu=1}^{N_a} s_{\nu,a} \tag{4.14}$$

da cui l'importante relazione

$$S_a = \sum_{\nu=1}^{N_a} s_{\nu,a} \tag{4.15}$$

S_a è la dimensione del guscio enrgetico $\mathcal{H}_a$, N_a è il numero di celle di fase su tale guscio, mentre $s_{\nu,a}$ è la dimensione di ciascuna cella di fase $\mathcal{H}_{\nu,a}$ su tale guscio, per a fissato.

Osservazioni.

1)Poichè il prodotto tra $\sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a}$ e $\sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,b}$ ($a \neq b$, gusci energetici differenti $\mathcal{H}_a \neq \mathcal{H}_b$) è uguale alla somma di quegli $\hat{\mathbb{E}}_p$ che appartengono ad entrambe le somme, e, d'altra parte, è uguale al prodotto di $\sum_{\rho=1}^{N_a} \hat{\mathbb{P}}_{\rho,a}$ e $\sum_{\rho=1}^{N_a} \hat{\mathbb{P}}_{\rho,b}$ che è nullo,segue che la somma degli $\hat{\mathbb{E}}_p$ comuni è zero. Dunque non ve ne sono, in quanto la somma di diversi $\hat{\mathbb{E}}_p$, cioè di diversi $\hat{\mathbb{P}}_{\omega_{\lambda,p}}$ (p fissato) non può essere nulla.
(Perchè, da $\hat{\mathbb{P}}_{\omega_{1,p}} + \hat{\mathbb{P}}_{\omega_{2,p}} + ... = 0$ si otterrebbe, ad esempio moltiplicando per $\hat{\mathbb{P}}_{\omega_{1,p}}$, $\hat{\mathbb{P}}^2_{\omega_{1,p}} = \hat{\mathbb{P}}_{\omega_{1,p}} = 0$, cosa che è impossibile).

2)Gli $\hat{\mathbb{E}}_{\nu,a}$ esauriscono tutti gli $\hat{\mathbb{E}}_p$. Infatti è sufficiente mostrare che vale

$$\sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} = \sum_{p=1}^{\infty} \hat{\mathbb{E}}_p \tag{4.16}$$

Infatti entrambi i membri sono uguali a $\hat{1}$, perchè

$$\sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} = \sum_{a=1}^{\infty} \sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\rho,a} = \hat{1} \tag{4.17}$$

in quanto i $\phi_{\rho,a}$ formano un sistema ortonormale completo. E

$$\sum_{p=1}^{\infty} \hat{\mathbb{E}}_p = \sum_{p=1}^{\infty} \sum_{\lambda=1}^{s_p} \hat{\mathbb{P}}_{\omega_{\lambda,p}} = \hat{1} \tag{4.18}$$

in quanto anche gli $\omega_{\lambda,p}$ formano un sistema ortonormale completo.

3) Ne consegue che $\hat{\mathbb{E}}_{\nu,a}$ e $s_{\nu,a}$ con $a = 1, 2, ...$ e $\nu = 1, 2, ..., N_a$ è solo una maniera differente di indicizzare $\hat{\mathbb{E}}_p$ e s_p , con p=1,2,...

4) Corrispondentemente von Neumann scrive $\omega_{\lambda,\nu,a}$ in luogo di $\omega_{\lambda,p}$, dove

$a = 1, 2, ...,\quad \nu = 1, 2, ..., N_a,\quad \lambda = 1, 2, ..., s_{\nu,a},$
e definisce l'operatore

$$\hat{\Delta}_a \stackrel{def}{=} \sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}} = \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} = \sum_{\nu=1}^{N_a} \sum_{\lambda=1}^{s_{\nu,a}} \hat{\mathbb{P}}_{\omega_{\lambda,\nu,a}} \tag{4.19}$$

5) Da tutto ciò si deduce che $\frac{1}{S_a}\hat{\Delta}_a$ è la miscela di stati $\phi_{1,a}, ..., \phi_{S_a,a}$ con pesi uguali, o, in alternativa, la miscela di miscele di $\frac{1}{s_{1,a}}\hat{\mathbb{E}}_{1,a}, ..., \frac{1}{s_{N_a,a}}\hat{\mathbb{E}}_{N_a,a}$ con pesi $s_{1,a}, ..., s_{N_a,a}$.

L'analogia con la meccanica statistica di Gibbs

E' ovvio ora che :
$\frac{1}{S_a}\hat{\Delta}_a$ corrisponde alla superficie di energia a-esima, $\mathcal{H}_a$
N_a è il numero di celle di fase $\hat{\mathbb{E}}_{\nu,a}$ sulla superficie di energia a-esima
la dimensione $S_a = Tr(\hat{\Delta}_a)$ rappresenta il numero di orbite quantistiche stazionarie, cioè il numero di microstati su $\mathcal{H}_a$

Verifica del fatto che $\frac{1}{S_a}\hat{\Delta}_a$ ha traccia uguale a 1.

Innanzitutto è

$$\frac{1}{S_a}\hat{\Delta}_a = \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} = \frac{1}{S_a} \sum_{\nu=1}^{N_a} \sum_{\lambda=1}^{s_{\nu,a}} \hat{\mathbb{P}}_{\omega_{\lambda,\nu,a}} =$$

$$= \frac{1}{S_a}\left(\sum_{\lambda=1}^{s_{1,a}} |\, \omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a} \,| + \sum_{\lambda=1}^{s_{2,a}} |\, \omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a} \,| + ... \right.$$

$$\left. + \sum_{\lambda=1}^{s_{N_a,a}} |\, \omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a} \,| \right) \tag{4.20}$$

da cui

$$Tr(\frac{1}{S_a}\hat{\Delta}_a) = \frac{1}{S_a}\left(\sum_{\lambda=1}^{s_{1,a}} Tr(|\,\omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a}\,|) + \sum_{\lambda=1}^{s_{2,a}} Tr(|\,\omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a}\,|) + ...\right.$$

$$\left. + \sum_{\lambda=1}^{s_{N_a,a}} Tr(|\,\omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a}\,|)\right) =$$

$$= \frac{1}{S_a}\left(s_{1,a} + s_{2,a} + ... + s_{N_a,a}\right) = \frac{1}{S_a}\sum_{\nu=1}^{N_a} s_{\nu,a} = \frac{S_a}{S_a} = 1 \qquad (4.21)$$

Osservazione

Le possibili misurazioni macroscopiche dell'energia decompongono la totalità dei possibili stati nelle superfici di energia $\hat{\Delta}_a$, per $a = 1, 2, ...$ Ulteriori misurazioni macroscopiche dell'energia, che risolverebbero la $\hat{\Delta}_a$ nelle $\phi_{\rho,a}$, per $\rho = 1, 2, ..., S_a$, non sono possibili.
Sono possibili tuttavia misurazioni di grandezze che non commutano con $\hat{H}$, ossia di non integrali del moto. Tali misurazioni decompongono la superficie di energia nelle celle di fase $\hat{\mathbb{E}}_{\nu,a}$, per $\nu = 1, 2, ...N_a$.
Ma un'ulteriore decomposizione delle $\hat{\mathbb{E}}_{\nu,a}$, che risolverebbe le $\hat{\mathbb{E}}_{\nu,a}$ nelle $\omega_{\lambda,\nu,a}$, $\lambda = 1, 2, ...s_{\nu,a}$, risulta macroscopicamente impossibile.
L'ordine di grandezza di $s_{\nu,a}$ è una misura dell'inaccuratezza dei metodi macroscopici di misurazione.
Mentre N_a (numero di celle di fase $\mathcal{H}_{\nu,a}$ sulla superficie di energia $\mathcal{H}_a$) è una misura di quanto sono adeguati i metodi di misurazione macroscopici di grandezze che variano nel tempo (i non-integrali del moto) e che quindi non sono misurabili simultaneamente all'energia.
Vale la relazione

$$S_a = \sum_{\nu=1}^{N_a} s_{\nu,a} \qquad (4.22)$$

dove:
S_a è la dimensione della superficie di energia $\hat{\mathcal{H}}_a$ (o il suo proiettore $\hat{\Delta}_a$,

nella terminologia di von Neumann): $S_a = din(\hat{\mathcal{H}}_a)$
$s_{\nu,a}$ è la dimensione della ν-esima cella di fase $\hat{\mathcal{H}}_{\nu,a}$ (o il suo proiettore $\hat{\mathcal{E}}_{\nu,a}$ nella terminologia di von Neumann): $s_{\nu,a} = dim(\hat{\mathcal{H}}_{\nu,a})$
N_a è il numero di celle di fase sulla a-esima superficie di energia.
Ordini di grandezza tipici sono
$N_a \sim 10^{20}$
$s_{\nu,a} \sim 10^{10^{20}}$
Ed è $S_a \gg s_{\nu,a}$
(Si vedrà in seguito che condizione necessaria per la validità dell'ergodicità è che $s_{\nu,a} \gg N_a$).

Probabilità di transizione e ensemble microcanonico.

Sia dato un arbitrario stato ψ normalizzato: $\|\psi\|^2 =< \psi \mid \psi >= 1$. Sia $\mathcal{A}$ un valore della generica osservabile $\hat{A}$, $\mathcal{E}$ un valore della osservabile energia $\hat{H}$ e sia $\mathcal{H}_{\nu,a}$ il sottospazio della ν-esima cella di fase $\hat{\mathbb{E}}_{\nu,a}$ sulla a-esima superficie di energia $\hat{\Delta}_a$, il cui sottospazio corrispondente sia $\mathcal{H}_a$:

$$\mathcal{H}_a = \bigoplus_{\nu=1}^{N_a} \mathcal{H}_{\nu,a} \tag{4.23}$$

La probabilità che una misurazione macroscopica di $\hat{A}$, sul sistema che si trova nello stato ψ, fornisca valori nella cella di fase $\mathcal{H}_{\nu,a}$, è data dalla somma delle probabilità di transizione di ψ alle autofunzioni $\omega_{1,\nu,a}, \omega_{2,\nu,a}, ..., \omega_{s_{\nu,a},\nu,a}$ che generano $\mathcal{H}_{\nu,a}$:

$$\mathbb{P}\Big\{\mathcal{A} \in \mathcal{H}_{\nu,a}\Big\} = \sum_{\lambda=1}^{s_{\nu,a}} | < \omega_{\lambda,\nu,a} \mid \psi > |^2 = \sum_{\lambda=1}^{s_{\nu,a}} < \hat{\mathbb{P}}_{\omega_{\lambda,\nu,a}} \psi \mid \psi >=< \hat{\mathbb{E}}_{\nu,a} \psi \mid \psi > \tag{4.24}$$

Questo numero ci dice quanto la cella $\mathcal{H}_{\nu,a}$ è occupata dallo stato ψ.
Analogamente, la probabilità che il valore di energia $\mathcal{E}$appartenga all'insieme $\{W_{1,a}, W_{2,a}, ..., W_{S_a,a}\}$ è:

$$\mathbb{P}\Big\{\mathcal{E} \in \{W_{1,a}, W_{2,a}, ..., W_{S_a,a}\}\Big\} = \sum_{\rho=1}^{S_a} | < \phi_{\rho,a} \mid \psi > |^2 = \sum_{\rho=1}^{S_a} < \hat{\mathbb{P}}_{\phi_{\rho,a}} \psi \mid \psi >=< \Delta_a \psi \mid \psi >$$

Questo numero ci dice quanto la superficie di energia $\mathcal{H}_a$ è occupata dallo stato ψ, ovvero è il numero di occupazione di $\mathcal{H}_a$
Da notare che si ha

$$1 =< \psi \mid \psi >= \sum_{a=1}^{\infty} < \hat{\Delta}_a \psi \mid \psi >= \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} < \hat{\mathbb{E}}_{\nu,a} \psi \mid \psi > \tag{4.25}$$

Pertanto von Neumann definisce l'operatore statistico dell'ensemble microcanonico quantistico $\hat{\mathbb{U}}_\psi$, relativo allo stato ψ, come mistura di $\frac{1}{S_1}\hat{\Delta}_1, \frac{1}{S_2}\hat{\Delta}_2, \ldots.$ con pesi rispettivi $< \hat{\Delta}_1 \psi \mid \psi >, < \hat{\Delta}_2 \psi \mid \psi >, \ldots.$:

$$\hat{\mathbb{U}}_\psi \stackrel{def}{=} \sum_{a=1}^{\infty} < \hat{\Delta}_a \psi \mid \psi > \frac{\hat{\Delta}_a}{S_a} - \tag{4.26}$$

E' facile vedere che tale operatore ha traccia uguale a 1.

4.2 Implementazione della Dimostrazione

Von Neumann definisce

$$\psi_0 \stackrel{def}{=} \sum_{a=1}^{\infty} \sum_{\rho=1}^{S_a} r_{\rho,a} r^{-i\alpha_{\rho,a}} \phi_{\rho,a} \tag{4.27}$$

con $r_{\rho,a} \geq 0$ e $0 \leq \alpha_{\rho,a} \leq 2\pi$
e, data l'equazione di Schroedinger $\partial_t \psi_t = \frac{-i}{h} \hat{H} \psi_t$,
definisce l'operatore energia come

$$\hat{H} \stackrel{def}{=} \sum_{a=1}^{\infty} \sum_{\rho=1}^{S_a} W_{\rho,a} \hat{\mathbb{P}}_{\phi_{\rho,a}} \tag{4.28}$$

(Da notare il legame

$$\hat{\Delta}_a = \sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}} = \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} = \sum_{\nu=1}^{N_a} \sum_{\lambda=1}^{s_{\nu,a}} \hat{\mathbb{P}}_{\omega_{\lambda,\nu,a}} \tag{4.29}$$

$$\tag{4.30}$$

dove è $\hat{A}\omega_{\lambda,\nu,a} \overset{def}{=} \eta_{\nu,a}\omega_{\lambda,\nu,a})$
Essendo $\psi_t = e^{-\frac{i}{h}W_{\rho,a}t}\psi_0$
si ha

$$\psi_t = \sum_{a=1}^{\infty}\sum_{\rho=1}^{S_a} r_{\rho,a} e^{-i(\frac{W_{\rho,a}}{h}+\alpha_{\rho,a})}\phi_{\rho,a} \tag{4.31}$$

Quindi, von Neumann introduce le seguenti abbreviazioni

$$x_{\nu,a} \overset{def}{=} < \hat{\mathbb{E}}_{\nu,a}\psi_t \mid \psi_t > \tag{4.32}$$

e

$$u_a \overset{def}{=} < \hat{\Delta}_a\psi_t \mid \psi_t > \equiv < \hat{\Delta}_a\psi_0 \mid \psi_0 > \tag{4.33}$$

dove $u_a =< \hat{\Delta}_a\psi_0 \mid \psi_0 >$ deriva dal fatto che

$$\begin{aligned} u_a \overset{def}{=} < \hat{\Delta}_a\psi_t \mid \psi_t > = < \sum_{\rho=1}^{S_a} \hat{\mathbb{P}}_{\phi_{\rho,a}}\psi_t \mid \psi_t > = \\ \sum_{\rho=1}^{S_a} < \psi_t \mid \phi_{\rho,a} >< \phi_{\rho,a} \mid \psi_t > = \sum_{\rho=1}^{S_a} \mid < \psi_t \mid \phi_{\rho,a} > \mid^2 = \\ = \sum_{\rho=1}^{S_a} r_{\rho,a}^2 = cost. \end{aligned} \tag{4.34}$$

indipendente da t.
Si può notare che si ha

$$\sum_{\nu=1}^{N_a} x_{\nu,a} = < \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a}\psi_t \mid \psi_t > = < \hat{\Delta}_a\psi_t \mid \psi_t > = u_a \tag{4.35}$$

e

$$\sum_{a=1}^{\infty} u_a =< \sum_{a=1}^{\infty} \Delta_a \psi_t \mid \psi_t >=< \hat{1}\psi_t \mid \psi_t >= 1 \tag{4.36}$$

Quindi

$$\sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} x_{\nu,a} = \sum_{a=1}^{\infty} u_a = 1. \tag{4.37}$$

4.2.1 Osservabili macroscopiche, medie nello stato ψ e nell'ensemble

Sia ora A una qualsiasi osservabile macroscopica, tale che

$$\hat{\mathbb{A}}\ \omega_{\lambda,\nu,a} \stackrel{def}{=} \eta_{\nu,a}\ \omega_{\lambda,\nu,a} \tag{4.38}$$

Allora

$$\hat{\mathbb{A}} = \sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} \eta_{\nu,a}\hat{\mathbb{E}}_{\nu,a} = \sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} \eta_{\nu,a} \sum_{\lambda=1}^{s_{\nu,a}} \mid \omega_{\lambda,\nu,a} >< \omega_{\lambda,\nu,a} \mid \tag{4.39}$$

Dunque, le $\omega_{\lambda,\nu,a}$ della cella di fase $\hat{\mathbb{E}}_{\nu,a}$, che ha dimensione $s_{\nu,a}$, sono autofunzioni di $\hat{\mathbb{A}}$ con autovalore $\eta_{\nu,a}$: $\eta_{\nu,a}$ è il valore di $\hat{\mathbb{A}}$ nella cella di fase $\hat{\mathbb{E}}_{\nu,a}$.
I valori di aspettazione, $\mathbb{E}_\psi(A)$ e $\mathbb{E}_{\hat{U}_\psi}(A)$, nello stato ψ_t e nell'ensemble microcanonico $\hat{U}_\psi$ rispettivamente, sono

$$\mathbb{E}_\psi(A) \stackrel{def}{=} < \mathbb{A}\psi_t \mid \psi_t >= \sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} \eta_{\nu,a} < \hat{\mathbb{E}}_{\nu,a}\psi_t \mid \psi_t >= \sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} \eta_{\nu,a} x_{\nu,a} \tag{4.40}$$

e

$$\mathbb{E}_{\hat{U}_\psi}(A) \stackrel{def}{=} = Tr(\hat{\mathbb{A}}\hat{U}_\psi) = Tr\left[\left(\sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} \eta_{\nu,a}\hat{\mathbb{E}}_{\nu,a}\right)\left(\sum_{b=1}^{\infty}\sum_{\nu=1}^{N_a} \frac{u_a}{S_a}\hat{\mathbb{E}}_{\nu,b}\right)\right] = $$
$$= \sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a} \eta_{\nu,a}\frac{s_{\nu,a}}{S_a} u_a$$

poichè $\hat{\mathbb{E}}_{\nu,a}\hat{\mathbb{E}}_{\nu,b} = 0$ per $a \neq b$ e $Tr(\hat{\mathbb{E}}^2_{\nu,a}) = Tr(\hat{\mathbb{E}}_{\nu,a}) = s_{\nu,a}$

4.2.2 Stima di $(\mathbb{E}_{\psi_t}(\hat{A}) - \mathbb{E}_{\hat{U}_\psi}(\hat{A}))^2$ mediante disuguaglianza di Schwarz.

Si può scrivere:

$$(\mathbb{E}_{\psi_t}(\hat{A}) - \mathbb{E}_{\hat{U}_\psi}(\hat{A}))^2 = \left(\sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \eta_{\nu,a} \left[x_{\nu,a} - \frac{s_{\nu,a} u_a}{S_a} \right] \right)^2 =$$

$$= \left(\sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \sqrt{\frac{s_{\nu,a} u_a}{S_a}} \eta_{\nu,a} \sqrt{\frac{S_a}{s_{\nu,a} u_a}} \left[x_{\nu,a} - \frac{s_{\nu,a} u_a}{S_a} \right] \right)^2 \leq$$

$$\leq \left(\sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{s_{\nu,a} u_a}{S_a} \eta^2_{\nu,a} \right) \left(\sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a} u_a} \left[s_{\nu,a} - \frac{s_{\nu,a} u_a}{S_a} \right]^2 \right) \tag{4.41}$$

per la disuguaglianza di Schwarz $\left((\sum_n a_n b_n)^2 \leq \sum_n a_n^2 \sum_n b_n^2 \right)$
Si pone il primo fattore uguale a $\overline{\eta^2}_{\nu,a}$

$$\overline{\eta^2}_{\nu,a} \stackrel{def}{=} \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{s_{\nu,a} u_a}{S_a} \eta^2_{\nu,a} \tag{4.42}$$

Poichè

$$\frac{s_{\nu,a} u_a}{S_a} \geq 0; \quad \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{s_{\nu,a} u_a}{S_a} = 1; \quad \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{s_{\nu,a} u_a}{S_a} \eta^2_{\nu,a} \stackrel{def}{=} \overline{\eta^2} \tag{4.43}$$

quest'ultima espressione è una media pesata degli autovalori $\eta^2_{\nu,a}$ di $\hat{\mathbb{A}}^2$.
In realtà è la media microcanonica di $\hat{\mathbb{A}}^2$. Infatti

$$\mathbb{E}_{\hat{U}_\psi}(\hat{\mathbb{A}}^2) = Tr(\hat{U}_\psi \mathbb{A}^2) = \sum_{a=1}^{\infty} \frac{u_a}{S_a} Tr(\hat{\Delta}_a \hat{\mathbb{A}}^2) = \sum_{a=1}^{\infty} \frac{u_a}{S_a} Tr\left(\sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} \sum_{b=1}^{\infty} \sum_{\nu=1}^{N_a} \eta^2_{\nu,b} \hat{\mathbb{E}}^2_{\nu,b} \right) =$$

$$= \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{u_a}{S_a} \eta^2_{\nu,a} Tr\left(\hat{\mathbb{E}}_{\nu,a} \right) = \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{u_a}{S_a} s_{\nu,a} \eta^2_{\nu,a} = \overline{\eta^2}$$

Dunque $\overline{\eta}$ è una ragionevole misura dell'ordine di grandezza della osservabile $\hat{\mathbb{A}}$.
Ora si ha

$$\left(\mathbb{E}_\psi(\hat{\mathbb{A}}) - \mathbb{E}_{\hat{U}_\psi}(\hat{\mathbb{A}}) \right)^2 \leq \overline{\eta^2} \sum_{a=1}^{\infty} \sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a} u_a} \left[x_{\nu,a} - \frac{s_{\nu,a} u_a}{S_a} \right]^2 \tag{4.44}$$

4.2.3 Media rispetto al tempo, $\mathfrak{M}_t$

$$\mathfrak{M}_t\left\{\left(\mathbb{E}_\psi(\hat{\mathbb{A}})-\mathbb{E}_{\hat{U}_\psi}(\hat{\mathbb{A}})\right)^2\right\}\leq\overline{\eta^2}\mathfrak{M}_t\left\{\sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a}\frac{S_a}{s_{\nu,a}u_a}\left[x_{\nu,a}-\frac{s_{\nu,a}u_a}{S_a}\right]^2\right\}\quad(4.45)$$

La tesi ora è

$$\mathfrak{M}_t\left\{\sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a}\frac{S_a}{s_{\nu,a}u_a}\left[x_{\nu,a}-\frac{s_{\nu,a}u_a}{S_a}\right]^2\right\}=\epsilon\qquad(4.46)$$

uniformemente per tutti gli ψ_0, cioè per tutti gli $r_{\rho,a}$,$\alpha_{\rho,a}$ tali che $\sum_{a=1}^{\infty}\sum_{\rho=1}^{S_a}r_{\rho,a}^2=\|\psi_0\|^2=1$.
Si noti che, mentre $x_{\nu,a}$ dipende da t, $r_{\rho,a}$ e $\alpha_{\rho,a}$, u_α dipende solo da $r_{\rho,a}$.
Calcolo degli $x_{\nu,a}$:

$$x_{\nu,a}\overset{def}{=}<\hat{\mathbb{E}}_{\nu,a}\psi_t\mid\psi_t>=<\sum_{b=1}^{\infty}\sum_{\rho=1}^{S_b}r_{\rho,b}e^{-i(W_{\rho,b}t/\hbar+\alpha_{\rho,b})}\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,b}\mid\sum_{a=1}^{\infty}\sum_{\sigma=1}^{S_a}r_{\sigma,a}e^{-i(W_{\sigma,a}t/\hbar+\alpha_{\sigma,a})}\phi_{\sigma,a}>=$$

$$=\sum_{\rho,\sigma=1}^{S_a}r_{\rho,a}r_{\sigma,a}e^{-i[(W_{\rho,a}-W_{\sigma,a})t/\hbar+(\alpha_{\rho,a}-\alpha_{\sigma,a})]}<\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\mid\phi_{\sigma,a}>$$

poichè $<\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,b}\mid\phi_{\sigma,c}>=0$ a meno che $a=b=c$.
Ora, usando il fatto che $u_a=\sum_{\rho=1}^{S_a}r_{\rho,a}^2$, si ha

$$x_{\nu,a}-\frac{s_{\nu,a}u_a}{S_a}=\sum_{\substack{\rho,\sigma=1\\ \rho\neq\sigma}}^{S_a}r_{\rho,a}r_{\sigma,a}e^{-i[(W_{\rho,a}-W_{\sigma,a})t/\hbar+(\alpha_{\rho.a}-\alpha_{\sigma,a})]}<\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\mid\phi_{\sigma,a}>+$$

$$+\sum_{\rho=1}^{S_a}r_{\rho,a}^2\left\{<\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\mid\phi_{\rho,a}>-\frac{s_{\nu,a}}{S_a}\right\}$$

Quadrando questa espressione e mediando rispetto a t, tutti i termini della forma e^{ict}, con $c\neq 0$ si annullano.
Quindi, se per $\rho\neq\sigma$ è $W_\rho-W_\sigma\neq 0$, per $\rho\neq\sigma$, $\rho'\neq\sigma'$, è $(W_\rho-W_\sigma)-((W_{\rho'}-W_{\sigma'})\neq 0$, a meno che $\rho=\rho'$ e $\sigma=\sigma'$, cioè se per ogni a fissato tutti i $W_{1,a},W_{2,a},\ldots$ sono distinti, e così pure i $W_{\rho,a}-W_{\sigma,a}$

per $\rho \neq \sigma,\ \rho, \sigma = 1, 2, ...$
allora si ottiene

$$\mathfrak{M}_t\left\{\left[x_{\nu,a} - \frac{s_{\nu,a}u_a}{S_a}\right]^2\right\} = \sum_{\substack{\rho,\sigma=1\\(\rho\neq\sigma)}}^{S_a} r^2_{\rho,a} r^2_{\sigma,a} \mid < \hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\sigma,a} > \mid^2 +$$

$$+\left(\sum_{\rho=1}^{S_a} r^2_{\rho,a}\left\{ < \hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a} > -\frac{s_{\nu,a}}{S_a}\right\}\right)^2 \quad (4.47)$$

Von Neumann pone:

$$\mathcal{M}_{\nu,a} \stackrel{def}{=} \max_{\substack{\rho,\sigma=1\\ \rho\neq\sigma}}^{S_a}\left(\mid < \hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\sigma,a} > \mid^2 \right) \quad (4.48)$$

$$\mathcal{N}_{\nu,a} \stackrel{def}{=} \max_{\rho=1}^{S_a}\left(\left\{ < \hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a} > -\frac{s_{\nu,a}}{S_a}\right\}^2\right) \quad (4.49)$$

dove $\mathcal{M}_{\nu,a}$ e $\mathcal{N}_{\nu,a}$ sono chiaramente delle costanti, indipendenti cioè da $t, r_{\rho,a}, \alpha_{\rho,a}$ e quindi da ψ_t.

Poichè $\sum_{\rho=1}^{S_a} r^2_{\rho,a} = u_a$, si ha che

$$\mathfrak{M}_t\left\{\left[x_{\nu,a} - \frac{s_{\nu,a}u_a}{S_a}\right]^2\right\} \leq \sum_{\substack{\rho,\sigma=1\\(\rho\neq\sigma)}}^{S_a} r^2_{\rho,a} r^2_{\sigma,a}\mathcal{M}_{\nu,a} + (\sum_{\rho=1}^{S_a} r^2_{\rho,a}\sqrt{\mathcal{N}_{\nu,a}})^2 =$$

$$= u_a^2\mathcal{M}_{\nu,a} + (u_a\sqrt{\mathcal{N}_{\nu,a}})^2 = u_a^2(\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \quad (4.50)$$

E così

$$\mathfrak{M}_t\left\{\sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a}\frac{S_a}{s_{\nu,a}u_a}\left[x_{\nu,a} - \frac{s_{\nu,a}u_a}{S_a}\right]^2\right\} \leq \sum_{a=1}^{\infty}\sum_{\nu=1}^{N_a}\frac{S_a u_a}{s_{\nu,a}}(\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \quad (4.51)$$

Poichè $\sum_{a=1}^{\infty} u_a = 1$, questa è

$$\leq max_{a=1,2,...}\sum_{\nu=1}^{N_a}\frac{S_a}{s_{\nu,a}}(\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \quad (4.52)$$

Sarà sufficiente considerare il max solo degli a tali che $u_a \neq 0$, cioè su quelle superfici di energia $\mathcal{H}_a$ che occorrono nell'ensemble microcanonico.
Il teorema sarà quindi dimostrato se si mostra che, per questi a, si ha

$$\sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}} (\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) = \epsilon \tag{4.53}$$

E' quindi necessario ora trovare delle maggiorazioni di $\mathcal{M}_{\nu,a}$ e $\mathcal{N}_{\nu,a}$ opportunamente piccole.

4.2.4 Maggiorazioni di $\mathcal{M}_{\nu,a}$ e $\mathcal{N}_{\nu,a}$ nei casi più sfavorevoli e nei casi più favorevoli

Riguardiamo $\hat{H}$, e quindi $W_{\rho,a}$ e $\phi_{\rho,a}$, come fissati, e così pure $N_a, S_a, s_{\nu,a}$ e $\hat{\Delta}_a$. Si varia semplicemente $\hat{\mathbb{E}}_{\nu,a}$ entro questi limiti; cioè si varia il sistema ortonormale $\omega_{\lambda,\nu,a}$ $(\nu = 1, ..., N_a; \lambda = 1, ...s_{\nu,a})$ soggetto alla condizione

$$\sum_{\nu=1}^{N_a} \sum_{\lambda=1}^{s_{\nu,a}} \hat{\mathbb{P}}_{\lambda,\nu,a} \stackrel{def}{=} \sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,a} = \hat{\Delta}_a \tag{4.54}$$

dove si è posto $\hat{\mathbb{E}}_{\nu,a} \stackrel{def}{=} \sum_{\lambda=1}^{s_{\nu,a}} \hat{\mathbb{P}}_{\lambda,\nu,a} \quad per \quad \nu = 1, ..., N_a$.
Tali sistemi ortonorrmali $\omega_{\lambda,\nu,a}$ derivano uno dall'altro $\overline{\omega}_{\lambda,\nu,a}$ mediante trasformazioni unitarie in $S_a = \sum_{\nu=1}^{N_a} s_{\nu,a}$ dimensioni (dove a è tenuto fisso). Allora $\mathcal{M}_{\nu,a}, \mathcal{N}_{\nu,a}$ dipendono solo da $\omega_{\lambda,\nu,a}$.
Non per ogni scelta degli $\omega_{\lambda,\nu,a}$ essi sono piccoli come sarebbe necessario. Se, per esempio, gli $\omega_{\lambda,\nu,a}$ coincidono con i $\phi_{\rho,a}$, $\rho = 1, ...S_a$,a fissato, allora $< \hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a} >$ assumerebbe per un certo ρ il valore 1, e quindi si avrebbe

$$\mathcal{N}_{\nu,a} \geq (1 - \frac{s_{\nu,a}}{S_a}) \geq \frac{1}{4} > \epsilon \tag{4.55}$$

e poichè è sempre $s_{\nu,a} \leq \frac{1}{2} S_a$, si avrebbe

$$\sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}} (\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \geq \sum_{\nu=1}^{N_a} 2(0 + \frac{1}{4}) = \frac{N_a}{2} > \epsilon \tag{4.56}$$

che è arbitrariamente grande per N_a grande.
Ma questo caso sfavorevole sorge dall'assunzione che $\omega_{\lambda,\nu,a}$ non assuma il suo esatto significato fisico, perchè qui $\hat{\mathbb{E}}_{\nu,a}$ avrebbe le stesse autofunzioni di $\hat{H}$ (e quindi commuterebbe con $\hat{H}$), cosa che ci si aspetta non essere affatto il caso. Questo comportamento è singolare ed eccezionale, e per la schiacciante maggioranza degli $\omega_{\lambda,\nu,a}$ si troverà in seguito l'esatto ordine di grandezza di $\mathcal{M}_{\nu,a}\ \mathcal{N}_{\nu,a}$.
D'altra parte, nel caso più favorevole (ed inesatto) si può, invece di mediare

$$\mathcal{M}_{\nu,a} \stackrel{def}{=} \max_{\substack{\rho,\sigma=1\\ \rho\neq\sigma}}^{S_a} \left(|<\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\sigma,a}>|^2 \right) \tag{4.57}$$

$$\mathcal{N}_{\nu,a} \stackrel{def}{=} \max_{\rho=1}^{S_a} \left(\left\{ <\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a}> - \frac{s_{\nu,a}}{S_a} \right\}^2 \right) \tag{4.58}$$

su tutti i possibili sistemi $\omega_{\lambda.\nu,a}$, tentare di mediare gli argomenti dei *max* e poi considerarne i massimi. Allora, von Neumann ottiene [in Appendice A.3,pag 32]

$$\mathfrak{M}\left\{ |<\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\sigma,a}>|^2 \right\} = \tfrac{s_{\nu,a}(S_a-s_{\nu,a})}{S_a({S_a}^2-1)} \simeq \tfrac{s_{\nu,a}}{S_a^2} \qquad (\rho\neq\sigma) \tag{4.59}$$

$$\mathfrak{M}\left\{ <\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a}> \right\} = \tfrac{s_{\nu,a}}{S_a} \tag{4.60}$$

$$\mathfrak{M}\left\{ <\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a}> - \tfrac{s_{\nu,a}}{S_a} \right\}^2 = \tfrac{s_{\nu,a}(S_a-s_{\nu,a})}{S_a^2(S_a+1)} \simeq \tfrac{s_{\nu,a}}{S_a^2} \tag{4.61}$$

essendo sempre,in pratica, $s_{\nu,a} \ll S_a$.
Per $\mathcal{M}_{\nu,a}$ e $\mathcal{N}_{\nu,a}$ dunque si ottiene, inserendo $\frac{s_{\nu,a}}{S_a^2}$, che

$$\sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}}(\mathcal{M}_{\nu,a}+\mathcal{N}_{\nu,a}) = 2\sum_{\nu=1}^{N_a}\frac{1}{S_a} = \frac{2N_a}{S_a} < \epsilon \tag{4.62}$$

se $\frac{N_a}{S_a} \ll 1$, ovvero se $\dfrac{\sum_{\nu=1}^{N_a} s_{\nu,a}}{N_a} = \dfrac{S_a}{N_a} \gg 1$.
In conclusione, le $s_{\nu,a}$ devono essere grandi in media aritmetica.

4.2.5 Maggiorazioni esatte di $\mathcal{M}_{\nu,a}$ e $\mathcal{N}_{\nu,a}$

. Nelle Appendici A.5 e A.4 rispettivamente von Neumann ricava (con lunghi e laboriosi calcoli che qui non si riportano) le maggiorazioni estte seguenti:

$$\mathcal{M}_{\nu,a} \leq \frac{\ln S_a}{S_a} \tag{4.63}$$

e

$$\mathcal{N}_{\nu,a} \leq \frac{9 s_{\nu,a} \ln S_a}{S_a^2} \tag{4.64}$$

Da cui

$$\begin{aligned}\sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}} (\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \leq \sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}} \left(\frac{\ln S_a}{S_a} + \frac{9 s_{\nu,a} \ln S_a}{S_a^2} \right) = \\ = \ln S_a \left(\sum_{\nu=1}^{N_a} \frac{1}{s_{\nu,a}} + \frac{9 N_a}{S_a} \right)\end{aligned} \tag{4.65}$$

Introducendo le medie aritmetica e armonica, $\overline{s}_a$ *e* $\overline{\overline{s}}_a$, di $s_{\nu,a}$ $(\nu = 1, ..., N_a)$

$$\overline{s}_a \stackrel{def}{=} \frac{1}{N_a} \sum_{\nu=1}^{N_a} s_{\nu,a} \tag{4.66}$$

$$\overline{\overline{s}}_a \stackrel{def}{=} \frac{1}{N_a} \sum_{\nu=1}^{N_a} \frac{1}{s_{\nu,a}} \tag{4.67}$$

si ha

$$\sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}} (\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \leq \ln S_a \left(\frac{N_a}{\overline{\overline{s}}_a} + 9 \frac{1}{\overline{s}_a} \right) \simeq \ln S_a \frac{N_a}{\overline{\overline{s}}_a} \tag{4.68}$$

perchè $N_a \gg 1$ e $\frac{1}{\overline{s}_a} \leq \frac{1}{\overline{\overline{s}}_a}$ (la superficie di energia $\mathcal{H}_a$ contiene molte celle di fase)
E dunque

$$\sum_{\nu=1}^{N_a} \frac{S_a}{s_{\nu,a}} (\mathcal{M}_{\nu,a} + \mathcal{N}_{\nu,a}) \leq \ln S_a \frac{N_a}{\overline{\overline{s}}_a} = \ln S_a \sum_{\nu=1}^{N_a} \frac{1}{s_{\nu,a}} \tag{4.69}$$

Resta da stabilire quando questa espressione è uguale ad un ϵ.
Di certo si deve avere $\overline{s}_a \gg \overline{\overline{s}}_a \gg N_a$ da cui $\ln \overline{s}_a \geq \ln N_a$ cosicchè si può

sostituire (essendo $S_a = \overline{s}_a N_a$) $\ln S_a = \ln \overline{s}_a + \ln N_a$ con $\ln \overline{s}_a$.
Quindi la condizione è

$$(\ln \overline{s}_a)\frac{N_a}{\overline{\overline{s}}_a} \ll 1 \tag{4.70}$$

ovvero

$$\sum_{\nu=1}^{N_a} \frac{1}{s_{\nu,a}} \ll \frac{1}{\ln \overline{s}_a} \tag{4.71}$$

Questa relazione è verificata quando è

$$s_{\nu,a} \gg N_a \tag{4.72}$$

cioè quando la dimensione della generica cella di fase è molto maggiore del numero di celle di fase sulle superficie di energia costante. Conclusione:

$$\mathfrak{M}_t\left\{\left(\mathbb{E}_{\psi_t}(A) - \mathbb{E}_{\hat{U}_\psi}(A)\right)^2\right\} \leq \overline{\eta^2}\sum_{\nu=1}^{N_a}\frac{1}{s_{\nu,a}}\ln(S_a) = \epsilon \tag{4.73}$$

quando $s_{\nu,a} \gg N_a$.
Con ciò la dimostrazione del teorema ergodico quantistico è conclusa.

4.3 Osservazione: la disuguaglianza di $\breve{C}$ebi$\breve{s}$ev sulla sfera unitaria

L'operatore statistico microcanonico definito da von Neumann è

$$\hat{\mathbb{U}}_\psi \stackrel{def}{=} \sum_{a=1}^{\infty} \frac{< \Delta_a \psi_0 \mid \psi_0 >}{S_a} \Delta_a$$

Se calcoliamo il valore medio microcanonico $\mathbb{E}_{m.c.}$ del macrostato $\hat{\mathbb{E}}_{\nu,a}$, otteniamo

$$\begin{aligned} \mathbb{E}_{m.c.}\big\{macrostato\ \hat{\mathbb{E}}_{\nu,a}\big\} &= Tr(\mathbb{U}_\psi \hat{\mathbb{E}}_{\nu,a}) = \\ = \frac{\sum_{b=1}^{\infty} < \Delta_b \psi_0 \mid \psi_0 >}{S_a} & Tr\big(\sum_{\nu=1}^{N_a} \hat{\mathbb{E}}_{\nu,b} \hat{\mathbb{E}}_{\nu,a}\big) = \\ &= \frac{1}{S_a} Tr(\hat{\mathbb{E}}_{\nu,a}) = \frac{s_{\nu,a}}{S_a} \end{aligned} \tag{4.74}$$

Nella Appendice A.3, von Neumann ricava le seguenti relazioni

$$\mathbb{E}\big(\|\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\|^2\big) = \mathbb{E}(< \hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a} \mid \phi_{\rho,a} >) = \frac{s_{\nu,a}}{S_a} \tag{4.75}$$

per quasi ogni ψ_0 appartenente alla sfera unitaria in S_a dimensioni.
E inoltre

$$Var\big(\|\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\|^2\big) \stackrel{def}{=} \mathbb{E}\Bigg(\|\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\|^2 - \frac{s_{\nu,a}}{S_a}\Bigg)^2 < \frac{1}{s_{\nu,a}}(\frac{s_{\nu,a}}{S_a})^2 \tag{4.76}$$

Ciò significa che, per quasi ogni ψ_0 sulla sfera unitaria, la disuguaglianza di $\breve{C}$ebi$\breve{s}$ev (che è una forma della legge dei grandi numeri) si legge così:

$$\mathbb{P}\Bigg\{\bar{|}\|\hat{\mathbb{E}}_{\nu,a}\phi_{\rho,a}\|^2 - \frac{s_{\nu,a}}{S_a}\bar{|}\geq\epsilon\Bigg\} \leq \frac{1}{\epsilon^2}\frac{s_{\nu,a}}{S_a^2} \tag{4.77}$$

dove, per ϵ piccolo la grandezza di $\frac{1}{\epsilon^2}$ è controbilanciata dalla piccolezza del fattore $\frac{s_{\nu,a}}{S_a^2}$ quando è, come è sempre $s_{\nu,a} \ll S_a$.

Bibliografia

[1] Khinchin, Mathematical foundations of Statistical Mechanics

[2] Gallavotti, Meccanica statistica.Trattatello

[3] J. von Neumann, Proof of the Ergodic Theorem and the H-Theorem in Quantum Mechanics (Translator Roderich Tumulka)