

TORTUGA
Publisher

Sulla possibilità di una trattazione matematica della meccanica statistica classica

Note sulla lettura di *Mathematical Foundations of statistical Mechanics* di A. I. Khinchin e di *Stir in stillness* di J. van Lith

Alberto Maggi

[219,915]

55 via Lopez, 57010 Guasticce (LI)

0586 984 980

Sommario

Prefazione	5
I I fondamenti della meccanica statistica	7
I.1 Geometria e cinematica nello spazio delle fasi	7
I.1.1 Spazio delle fasi	7
I.1.2 Il teorema di Liouville	8
I.1.3 Il teorema di Birkhoff	10
I.1.4 Indecomponibilità metrica	14
I.1.5 Funzioni di struttura	17
I.1.6 Sistemi composti	18
I.2 Il problema ergodico	19
I.2.1 Interpretazione delle quantità fisiche in meccanica statistica	19
I.2.2 Integrali del moto	20
I.2.3 Indecomponibilità metrica delle varietà ridotte	21
I.2.4 Formulazione senza uso di indecomponibilità metrica	22
II Teoria delle probabilità e meccanica statistica	25
II.1 Riduzione del problema alla teoria delle probabilità	25
II.1.1 Legge fondamentale di distribuzione	25
II.1.2 Legge di distribuzione per una componente del sistema	25
II.1.3 Funzioni generatrici	27
II.1.4 Leggi di distribuzione coniugate	28
II.2 Applicazioni del teorema del limite centrale	29
II.2.1 Il teorema del limite centrale	30
II.2.2 La legge di Boltzmann	32
II.2.3 Valori medi delle funzioni additive	33
II.2.4 Legge di distribuzione per l'energia di una componente grande	36
II.2.5 Gas monoatomico ideale	37

II.2.6	Il teorema dell'equipartizione dell'energia	38
II.3	L'insieme canonico di Gibbs	40
II.3.1	Sistemi isolati e sistemi in equilibrio termico	40
II.3.2	Distribuzione canonica	41
III	I metodi della meccanica statistica	43
III.1	Impostazione del problema	43
III.2	L'approccio ergodico	44
III.3	L'approccio di Khinchin	45
III.4	La strategia di Malament, Zabell e Vranas	47

Prefazione

In questo testo vediamo come è possibile dare una fondazione matematica coerente (non priva di qualche piccolo intoppo) della meccanica statistica. Dal momento che il problema è in molte sue parti aperto ancora oggi, si vedrà che non tutte le assunzioni discendono direttamente dalla meccanica microscopica e dalla teoria statistica. Per smussare gli angoli delle teorie moderne di Malament, Zabell e Vranas c'è ancora un po' da lavorare! Tuttavia, sono certo che lo studio di quanto qui presentato, risulterà in una buona giustificazione (almeno in prima approssimazione) dell'altrimenti misterioso successo della strategia di Gibbs: il *phase averaging method*. Il testo non è del tutto organico, perché è costituito dagli appunti da me presi durante la lettura del Khinchin (il cui riassunto è la parte più consistente di quanto scritto) come della dissertazione di Janneke van Lith (veramente molto interessante) sulla storia della teoria della meccanica statistica, da Birkhoff fino a Vranas (ma direi pure fino alla van Lith, visti i suoi importanti contributi, specie per quanto concerne la nozione di equilibrio).

Questi appunti nascono dal mio desiderio, a tratti irresistibile, di studiare un po' più in dettaglio le connessioni tra gli *ensembles* di Gibbs e i postulati della fisica classica. Già, perché la presentazione istituzionale della meccanica statistica, spiegata con le vecchissime teorie di Boltzmann e Gibbs, faticanti dal punto di vista matematico, e molto lontane dalla fisica newtoniana (si pensi all'ipotesi ergodica di Boltzmann, così interessante e così palesemente falsa), mi è risultata a tal punto insopportabile, che ho deciso di andare a fondo al problema, per quanto mi era consentito dai vincoli del poco tempo a disposizione (in fin dei conti, nessuno mi chiederà mai queste cose, a nessun esame), e delle mie poche competenze in materia...

Nell'ultima parte mancano tutte le dimostrazioni e anche le definizioni non sono curate bene (questo proprio perché alla fine non avevo veramente più tempo), ma mi è bastato leggere la tesi di Janneke van Lith e il fatto che alcune dimostrazioni esistessero (e non molte attendessero poi di essere trovate!) per soddisfare la mia famelica curiosità in merito alla questione dei fondamenti di una delle teorie che ritengo più prolifiche e belle dell'intera fisica.

Guasticce, 15 ottobre 2001

Alberto Maggi

I fondamenti della meccanica statistica

Il teorema di Liouville e quello di Birkhoff stanno a fondamento della meccanica statistica. Il primo asserisce l'invarianza della misura di Lebesgue nell'evoluzione hamiltoniana di un sistema autonomo, il secondo, sfruttando pesantemente l'ipotesi di invarianza della misura, garantisce l'esistenza delle medie sul tempo delle quantità osservabili e con ciò la consistenza a livello matematico della misura fisica sul sistema.

In questo capitolo dimostriamo entrambi i teoremi (prima sezione) e presentiamo un'introduzione al problema ergodico (seconda sezione). In questo modo poniamo la fondazione della meccanica statistica in maniera critica e costruttiva. La dimostrazione del teorema di Birkhoff è fatta senza accennare minimamente alle ipotesi di ergodicità sulla linea di Kolmogoroff ripresa dal testo di riferimento per questi appunti: *Mathematical Foundations of Statistical Mechanics* traduzione dal russo del testo di **A. Khinchin**.

I.1 Geometria e cinematica nello spazio delle fasi

I.1.1 Spazio delle fasi

Spazio delle fasi Γ

Lo stato di un sistema meccanico (statistico) a s gradi di libertà può essere descritto assegnando i valori nel tempo delle coordinate canoniche $\mathbf{q} \equiv (q_1, \dots, q_s)$; $\mathbf{p} \equiv (p_1, \dots, p_s)$. Sia Γ lo spazio $2s$ -dimensionale i cui punti sono determinati dalle coordinate \mathbf{q}, \mathbf{p} . Ad ogni stato del sistema meccanico, che chiameremo brevemente G , corrisponde un ben determinato punto di Γ , **immagine** del sistema. L'intero spazio Γ si definisce **spazio delle fasi**.

Equazioni canoniche

Le equazioni del moto per le variabili canoniche sono date da, $i \in J_s$,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}; \quad (\text{I.1})$$

dove H (funzione delle variabili canoniche) è l'hamiltoniana del sistema meccanico. Nell'ipotesi, che assumiamo senz'altro, in cui H non dipenda dal tempo, si ha che H è un integrale primo del sistema dinamico (I.1), infatti,

$$\dot{H} = \frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial q_i} \dot{q}_i = -\frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} = 0$$

dove si è usata la convenzione di Einstein sugli indici ripetuti.

Omogeneità del tempo, traiettoria nello spazio delle fasi

Infine, l'orbita soluzione del sistema dinamico hamiltoniano è completamente determinata una volta che siano assegnati i dati iniziali, visto che il sistema è del primo ordine. Ne viene l'evoluzione di uno stato del sistema G al tempo t_0 , è nota univocamente, una volta noto lo stato al tempo t_0 . Dal punto di vista geometrico, l'insieme dei punti toccati dal punto immagine nello spazio Γ , cioè la **traiettoria** (od orbita), è univocamente determinata da un qualsiasi suo punto, cioè a dire, da ogni punto di Γ passa una e una sola traiettoria (questo grazie all'omogeneità del tempo che rende possibile la traslazione temporale del problema di Cauchy). Dal punto di vista operatoriale, fissati gli istanti t_1, t_2 , il sistema dinamico hamiltoniano che descrive il cambiamento nel tempo dello stato del sistema G , definisce una corrispondenza biunivoca nello spazio delle fasi che a ogni punto $M_1 \in \Gamma$ associa il punto M_2 evoluto al tempo t_2 del punto M_1 considerato come condizione iniziale assunta al tempo t_1 . L'operatore in questione, $U(t_1, t_2)$, si dice **operatore di evoluzione temporale**. L'omogeneità del tempo (che si traduce nell'invarianza per traslazione temporale e nel fatto

che l'hamiltoniana è un integrale del moto) garantisce che la dipendenza di M_2 da M_1 è fissata una volta data la differenza $\Delta t \equiv t_2 - t_1$, sicché $U = U(t_2 - t_1) \equiv U(\Delta t)$, la qualcosa comporta che l'insieme degli operatori temporali al variare di Δt è un gruppo a un parametro.

Terminologia

In quanto segue capiterà di riferirci alle \mathbf{q}, \mathbf{p} come coordinate dinamiche del punto immagine e di chiamare ogni funzione delle \mathbf{q}, \mathbf{p} come funzione di fase. Ovviamente, la funzione di fase più importante per un sistema è la hamiltoniana.

Chiameremo invariante un sottoinsieme dello spazio delle fasi mappato in sé dall'evoluzione temporale a ogni istante.

Prima di chiudere ancora qualche nota sul sistema (I.1). Denotiamo con $\mathbf{x} \equiv (\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2s}$ allora la (I.1) diviene

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$$

dove il campo stazionario di velocità \mathbf{v} è dato da

$$\mathbf{v} = J(\nabla_{\mathbf{x}} H), \quad J \equiv \left(\begin{array}{c|c} 0 & \mathbb{I}_s \\ \hline -\mathbb{I}_s & 0 \end{array} \right).$$

Più in generale un problema del tipo

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{v}(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}$$

con $(t, \mathbf{x}) \in I \times \Omega$ (I intervallo e Ω aperto) dà luogo a un operatore di evoluzione $U(t, t_0)$ che a \mathbf{x}_0 associa $\mathbf{x}(t)$. Il teorema di differenziabilità (vedi **A. Maggi**, *Analisi II per Fisici*) asserisce che qualora $\mathbf{v} \in \mathcal{C}^1(I \times \Omega)$, allora $U(t, t_0) \in \mathcal{C}^1(I \times \Omega)$ (dove a variare è t e non t_0 , beninteso). Si ha inoltre, che $X(t) \equiv \partial U(t, t_0) / \partial \mathbf{y}$ risolve il sistema dinamico

$$\begin{cases} \dot{X} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}}(t, U(t, t_0) \mathbf{x}_0) X(t) \\ X(t_0) = \mathbb{I} \end{cases} \quad (\text{I.2})$$

Noi assumeremo sempre $H \in \mathcal{C}^2$, di modo che il nostro operatore di evoluzione temporale rispetti le ipotesi del teorema di differenziabilità del flusso.

La fondazione della meccanica statistica risiede in due teoremi legati alla struttura delle equazioni canoniche, il teorema di Liouville e quello di Birkhoff. Nelle prossime due sottosezioni andremo a dimostrarli.

1.1.2 Il teorema di Liouville

Evoluzione di insiemi misurabili: teorema di Liouville

Sia M un insieme Lebesgue misurabile di punti nello spazio delle fasi Γ di un dato sistema meccanico. Se assumiamo i punti di M come dati di Cauchy per l'evoluzione temporale in un tempo t , otteniamo che l'insieme M è mappato da $U(t)$ in un nuovo insieme $M_t \subset \Gamma$. Poiché l'operatore $U(t)$ è continuo, di modo che trasforma misurabili in misurabili, si ha che M_t è misurabile. Il teorema di Liouville sostiene che M_t ha la stessa misura di M . In altre parole, la misura di un qualsiasi insieme misurabile nello spazio delle fasi è un integrale del moto.

Dimostrazione del teorema di Liouville

Abbiamo

$$m(M_t) = \int_{M_t} dm(\mathbf{x})$$

cambiando variabile, essendo $U(t)$ continuo e invertibile (cioè un diffeomorfismo),

$$m(M_t) = \int_M \left| \det \frac{\partial U(t)}{\partial \mathbf{x}} \right| dm(\mathbf{x})$$

Si tratta dunque di vedere come evolve nel tempo il determinante jacobiano dell'operatore di evoluzione. A questo scopo dobbiamo utilizzare la (I.2). Dunque, per $t = 0$,

$$\frac{\partial U(0)}{\partial \mathbf{x}} = \mathbb{I} \implies \det \frac{\partial U(0)}{\partial \mathbf{x}} = 1$$

se mostriamo che a ogni istante

$$\frac{d}{dt} \det \frac{\partial U(t)}{\partial \mathbf{x}} = 0,$$

allora

$$\det \frac{\partial U(t)}{\partial \mathbf{x}} = 1 \implies m(M_t) = m(M).$$

Primo passo A questo scopo cominciamo con il mostrare che

$$\det(\mathbb{I} + \varepsilon A) = 1 + \varepsilon \operatorname{Tr} A + O(\varepsilon^2)$$

Sia $\lambda \in \mathbb{C}$ autovalore di A e sia $\mathbf{w} \in E(\lambda, A)$, allora

$$(\mathbb{I} + \varepsilon A) \mathbf{w} = (1 + \varepsilon \lambda) \mathbf{w}$$

dunque se $\{\lambda_i\}$ è l'insieme degli autovalori di A , concludiamo

$$\det(\mathbb{I} + \varepsilon A) = \prod_i (1 + \varepsilon \lambda_i) = 1 + \varepsilon \sum_i \lambda_i + O(\varepsilon^2) = 1 + \varepsilon \operatorname{Tr} A + O(\varepsilon^2)$$

Poiché, poi

$$\det(B + o(\varepsilon)) = \det B + o(\varepsilon)$$

si conclude

$$\det \left(\frac{\partial U(t)}{\partial \mathbf{x}} \right) = \det \left(\mathbb{I} + \frac{\partial \mathbf{v}(\mathbf{x}_0)}{\partial \mathbf{x}} t + o(t) \right) = 1 + t \operatorname{Tr} \frac{\partial \mathbf{v}(\mathbf{x}_0)}{\partial \mathbf{x}} + o(t)$$

sicché

$$\frac{d}{dt} \det \left(\frac{\partial U(t)}{\partial \mathbf{x}} \right) \Big|_{t=0} = \operatorname{Tr} \frac{\partial \mathbf{v}(\mathbf{x}_0)}{\partial \mathbf{x}}$$

Secondo passo D'altra parte

$$U(t) = U(t - t_1) \circ U(t_1)$$

perciò dalla *chain rule*

$$\frac{\partial U(t)}{\partial \mathbf{x}} = \frac{\partial U(t - t_1)}{\partial \mathbf{x}} \frac{\partial U(t_1)}{\partial \mathbf{x}}$$

dal teorema di Binet

$$\det \frac{\partial U(t)}{\partial \mathbf{x}} = \det \frac{\partial U(t - t_1)}{\partial \mathbf{x}} \det \frac{\partial U(t_1)}{\partial \mathbf{x}}$$

derivando rispetto a t e considerando il risultato all'istante $t = t_1$

$$\frac{d}{dt} \det \frac{\partial U(t)}{\partial \mathbf{x}} \Big|_{t=t_1} = \frac{d}{ds} \det \frac{\partial U(t - t_1)}{\partial \mathbf{x}} \Big|_{s=0} \det \frac{\partial U(t_1)}{\partial \mathbf{x}} = \operatorname{Tr} \frac{\partial \mathbf{v}(\mathbf{x}_0)}{\partial \mathbf{x}} \det \frac{\partial U(t_1)}{\partial \mathbf{x}}$$

cioè, in modo più compatto

$$\begin{cases} \frac{d}{dt} \det \frac{\partial U(t)}{\partial \mathbf{x}} = \operatorname{div} \mathbf{v}(\mathbf{x}_0) \det \frac{\partial U(t)}{\partial \mathbf{x}} \\ \det \frac{\partial U(0)}{\partial \mathbf{x}} = 1 \end{cases}$$

Conclusione Dal momento che il flusso hamiltoniano $\mathbf{v} = J(\nabla_{\mathbf{x}} H)$ ha evidentemente divergenza nulla, si ha il teorema di Liouville, cioè

$$\det \frac{\partial U(t)}{\partial \mathbf{x}} = 1.$$

Sia ora f una funzione di fase Lebesgue sommabile. Sia M un insieme misurabile di misura finita. Dal teorema di Liouville otteniamo

$$\int_{M_t} f(\mathbf{x}) dm(\mathbf{x}) = \int_M f(U(t)\mathbf{x}) dm(\mathbf{x}).$$

In particolare, se M è invariante

$$\int_M f(\mathbf{x}) dm(\mathbf{x}) = \int_M f(U(t)\mathbf{x}) dm(\mathbf{x}).$$

Riassumiamo quanto ottenuto

Teorema I.1
(di Liouville)

In un sistema hamiltoniano omogeneo governato dall'operatore di evoluzione temporale $U(t)$ la misura di Lebesgue di un sottoinsieme misurabile dello spazio delle fasi è U -invariante. In simboli, se $X \subset \Gamma$ e $X(t) = U(t)X$ immagine di X secondo $U(t)$, allora $X(t)$ è misurabile e risulta

$$m(X(t)) = m(X).$$

I.1.3 Il teorema di Birkhoff

**Asserto del
teorema di
Birkhoff**

Sia X un sottoinsieme di Γ che sia **invariante** e di **misura finita**. Inoltre, sia f una funzione di fase sommabile su X . Il teorema di Birkhoff afferma semplicemente l'esistenza del limite

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(U(t)\mathbf{x}) dt$$

per quasi ogni \mathbf{x} in X . Tale limite può essere interpretato come la media di f sulla traiettoria passante per \mathbf{x} , cioè come risultato della misura di f sul sistema quando si sappia che questo abbia assunto a un certo istante lo stato \mathbf{x} .

**Dimostrazione
in più passi**

Cominciamo con il notare che l'integrale ha senso. Infatti, f e U sono misurabili, perciò $t \mapsto f(U(t)\mathbf{x})$ è misurabile come composizione di misurabili. Detto questo, veniamo alla dimostrazione del teorema (sulla linea di Kolmogoroff). Avvertiamo che si tratta di una dimostrazione complicata e che verrà svolta in più passi.

Un po' di notazioni. Se n è un intero

$$\int_n^{n+1} f(U(t)\mathbf{x}) dt \equiv x_n(\mathbf{x})$$

$$\int_n^{n+1} |f(U(t)\mathbf{x})| dt \equiv y_n(\mathbf{x})$$

Notiamo, inoltre, che l'integranda, per composizione, è sommabile nello spazio di misura prodotto (t, \mathbf{x}) , visto che U è continua su tale spazio e f è sommabile in \mathbf{x} . Perciò, dal teorema di Tonelli, concludiamo che x_n e y_n sono sommabili in \mathbf{x} .

Lemma I.2 *Quasi ovunque in X ,*

$$\lim_{n \rightarrow \infty} \frac{y_n(\mathbf{x})}{n} = 0.$$

Dimostrazione Introduciamo la variabile $\tau = t - n$, allora

$$y_n(\mathbf{x}) = \int_0^1 |f(U(\tau+n)\mathbf{x})| d\tau = \int_0^1 |f(U(\tau)U(n)\mathbf{x})| d\tau = y_0(U(n)\mathbf{x}).$$

Definiamo adesso i sottoinsiemi di X , $E_{n,n}$ e $E_{n,0}$, formati rispettivamente dai punti \mathbf{x} tali che

$$y_n(\mathbf{x}) > n\varepsilon$$

$$y_0(\mathbf{x}) > n\varepsilon$$

dove $\varepsilon > 0$ è fissato. Ora, poiché

$$y_n(\mathbf{x}) = y_0(U(n)\mathbf{x})$$

si ha che se $\mathbf{x} \in E_{n,n}$, allora $U(n)\mathbf{x} \in E_{n,0}$, perciò dopo un tempo n i punti di $E_{n,n}$ vanno in $E_{n,0}$. Viceversa, se $\mathbf{y} \in E_{n,0}$, allora esiste un punto di $E_{n,n}$, $\mathbf{x} = U(-n)\mathbf{y}$, che dopo un tempo n va in \mathbf{y} :

$$y_n(\mathbf{x}) = y_0(\mathbf{y}) > n\varepsilon$$

Ne segue che

$$U(n)E_{n,n} = E_{n,0}$$

dunque, dal teorema di Liouville

$$m(E_{n,n}) = m(E_{n,0}).$$

Dove abbiamo sfruttato il fatto che $E_{n,n}$ e $E_{n,0}$ sono misurabili per costruzione essendo b_n una funzione misurabile.

Mostriamo adesso che la serie

$$\sum_{n=1}^{\infty} m(E_{n,n}) < \infty$$

Subito si ha che

$$\sum_{n=1}^{\infty} m(E_{n,n}) = \sum_{n=1}^{\infty} m(E_{n,0})$$

Se denotiamo con F_k l'insieme dei punti di X per cui

$$k\varepsilon < y_0(\mathbf{x}) \leq (k+1)\varepsilon$$

abbiamo

$$E_{n,0} = \bigcup_{k=n}^{+\infty} F_k$$

Gli insiemi F_k sono, ancora, misurabili e in più sono disgiunti, perciò

$$\begin{aligned} \sum_{n=1}^{\infty} m(E_{n,0}) &= \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} m(F_k) = \sum_{k=1}^{\infty} km(F_k) = \frac{1}{\varepsilon} \sum_{k=1}^{\infty} k\varepsilon m(F_k) = \\ &= \frac{1}{\varepsilon} \sum_{k=1}^{\infty} k\varepsilon \int_{k\varepsilon < b_0(\mathbf{x}) \leq (k+1)\varepsilon} dm(\mathbf{x}) = \frac{1}{\varepsilon} \sum_{k=1}^{\infty} \int_{k\varepsilon < b_0(\mathbf{x}) \leq (k+1)\varepsilon} k\varepsilon dm(\mathbf{x}) \leq \\ &\leq \frac{1}{\varepsilon} \int_X b_0(\mathbf{x}) dm(\mathbf{x}) = \frac{1}{\varepsilon} \int_X dm(\mathbf{x}) \int_0^1 d\tau |f(U(\tau)\mathbf{x})| = \\ &= \frac{1}{\varepsilon} \int_0^1 d\tau \int_X dm(\mathbf{x}) |f(U(\tau)\mathbf{x})| \end{aligned}$$

Dal momento che X è invariante

$$\int_X dm(\mathbf{x}) |f(U(\tau)\mathbf{x})| = \int_X dm(\mathbf{x}) |f(\mathbf{x})|$$

e siccome f è sommabile, ne viene che la serie è convergente. Questo significa che ogni punto di X , eccetto al più un insieme di misura nulla, non appartiene che a un numero finito di insiemi $E_{n,n}$ ($n \geq 1$), quindi per ogni $\mathbf{x} \in X$ (eccetto un insieme di misura nulla) esiste un intero $\nu(\mathbf{x})$ talché per ogni $n > \nu(\mathbf{x})$, risulta

$$y_n(\mathbf{x}) \notin E_{nn} \implies y_n(\mathbf{x}) \leq \varepsilon n$$

(c.v.d.) la tesi.

Osservazione I.1 Riguardiamo un attimo l'ultimo passaggio che non è proprio banale. Chiamiamo M l'insieme dei punti di X che appartengono a un numero infinito di $E_{n,n}$. Evidentemente,

$$\mathbf{x} \in M \iff \limsup_{n \rightarrow \infty} \chi_{E_{n,n}} = 1$$

Da cui

$$M = \bigcap_{n \in \mathbb{N}} \bigcup_{k > n} E_{k,k}$$

M è dunque dato dall'intersezione di misurabili incapsulati, essendo, se $n_1 < n_2$,

$$\bigcup_{k > n_2} E_{k,k} \subset \bigcup_{k > n_1} E_{k,k}$$

ne viene

$$m(M) = \lim_{n \rightarrow +\infty} m\left(\bigcup_{k > n} E_{k,k}\right) \leq \lim_{n \rightarrow +\infty} \sum_{k > n} m(E_{k,k}) = 0$$

- poiché la serie è convergente.

Poniamo, per ogni $a < b$,

$$h_{ab}(\mathbf{x}) \equiv \frac{1}{b-a} \int_a^b f(U(t)\mathbf{x}) dt,$$

Per definizione, se a e b fossero interi,

$$h_{ab}(\mathbf{x}) = \frac{1}{b-a} (x_a(\mathbf{x}) + x_{a+1}(\mathbf{x}) + \dots + x_{b-1}(\mathbf{x})).$$

Lemma I.3 Se $h_{0n}(\mathbf{x})$, con n intero, non ammette limite per $n \rightarrow \infty$ su un insieme M di misura positiva, allora esistono due reali $\alpha < \beta$ e un sottoinsieme M^* di M , tali che $m(M^*) > 0$ e per ogni $\mathbf{x} \in M^*$, risulta

$$l(\mathbf{x}) \equiv \liminf_{n \rightarrow \infty} h_{0n}(\mathbf{x}) < \alpha$$

$$L(\mathbf{x}) \equiv \limsup_{n \rightarrow \infty} h_{0n}(\mathbf{x}) > \beta$$

Dimostrazione

Consideriamo tutti gli intervalli $\delta_n = (\alpha_n, \beta_n)$ con estremi razionali. Se $\mathbf{x} \in M$ allora

$$l(\mathbf{x}) < L(\mathbf{x})$$

almeno per tutti i punti per cui h_{0n} non diverge, ma l'insieme dei punti in cui si ha divergenza ha misura nulla, essendo h_{0n} sommabile. Di conseguenza, tra tutti gli intervalli δ_n se ne trova uno per cui

$$l(\mathbf{x}) < \alpha_k < \beta_k < L(\mathbf{x})$$

Chiamiamo M_k l'insieme dei punti $\mathbf{x} \in M$ che sono associati a δ_k dalla disequazione di sopra. Chiaramente

$$M = \bigcup_k M_k$$

(c.v.d.) Siccome $m(M) > 0$ deve esistere almeno un $M_{\bar{k}}$ a misura non nulla. Posto $M^* \equiv M_{\bar{k}}$ e $\alpha \equiv \alpha_{\bar{k}}, \beta \equiv \beta_{\bar{k}}$, proviamo il lemma.

Assumiamo che le ipotesi del lemma siano soddisfatte. Sia $\mathbf{x} \in M^*$ e consideriamo l'intervallo (a, b) con $a < b$ interi. Chiamiamo l'intervallo **segmento proprio** del punto \mathbf{x} se sono rispettate le condizioni seguenti

$$\begin{aligned} h_{ab}(\mathbf{x}) &> \beta \\ h_{ab'}(\mathbf{x}) &\leq \beta, \quad a < b' < b \end{aligned}$$

Mostriamo che due segmenti propri (a_1, b_1) e (a_2, b_2) di uno stesso punto \mathbf{x} non possono sovrapporsi parzialmente l'uno all'altro. Infatti, se avessimo per esempio $a_1 < a_2 < b_1 < b_2$, allora avremmo

$$(b_1 - a_1) h_{a_1 b_1} = (a_2 - a_1) h_{a_1 a_2} + (b_1 - a_2) h_{a_2 b_1}$$

tuttavia

$$(b_1 - a_1) h_{a_1 b_1} > (b_1 - a_1) \beta$$

perciò

$$(b_1 - a_1) \beta < (a_2 - a_1) h_{a_1 a_2} + (b_1 - a_2) h_{a_2 b_1} \leq (a_2 - a_1) \beta + (b_1 - a_2) \beta = (b_1 - a_1) \beta$$

assurdo.

D'ora in poi chiameremo segmento proprio **massimale** di \mathbf{x} di rango s , un segmento proprio di lunghezza non superiore a s e non contenuto in un qualsiasi altro segmento proprio di lunghezza non superiore a s . Si vede subito che ogni segmento proprio di lunghezza non superiore a s è contenuto in uno e un solo segmento massimale di rango s . Infatti, tra tutti i segmenti propri di lunghezza inferiore a s ve ne sarà uno di lunghezza massima (il massimo esiste perché si tratta di numeri interi) e questo risulta massimale (si è dimostrata l'esistenza). Per quanto concerne l'unicità possiamo provarla per assurdo. Se esistessero due segmenti massimali essi dovrebbero avere il segmento di partenza in comune, perciò dovrebbero essere contenuti l'uno nell'altro, dunque uno dei due non sarebbe massimale.

Per ogni intero s denotiamo con M_s l'insieme dei punti \mathbf{x} di M^* per cui la disequaglianza

$$h_{0n}(\mathbf{x}) > \beta$$

è valida per almeno un $n \leq s$. Abbiamo allora

$$M^* = \bigcap_{s=1}^{+\infty} \bigcup_{k>s} M_k$$

infatti,

$$\begin{aligned} \mathbf{x} \in M^* &\iff \limsup_{s \rightarrow \infty} h_{0s}(\mathbf{x}) > \beta \iff \forall s \exists k > s : h_{0k}(\mathbf{x}) > \beta \iff \forall s \exists k > s : \mathbf{x} \in M_k \\ &\iff \mathbf{x} \in \bigcap_{s=1}^{+\infty} \bigcup_{k>s} M_k. \end{aligned}$$

Ne viene che, essendo $M_k \subset M_{k+1}$,

$$m(M^*) = \lim_{s \rightarrow \infty} m \left(\bigcup_{k>s} M_k \right) = \lim_{k \rightarrow \infty} m(M_k)$$

Per il teorema di permanenza del segno, essendo $m(M^*) > 0$, si conclude che per s sufficientemente grande $m(M_s) > 0$.

In quanto segue s sarà un intero per cui $m(M_s) > 0$.

Lemma I.4 *Condizione necessaria e sufficiente affinché $\mathbf{x} \in M_s$ è che \mathbf{x} ammetta un segmento proprio massimale (a, b) di rango s talché $a \leq 0 < b$.*

Dimostrazione Sia $\mathbf{x} \in M_s$ e sia n il più piccolo intero per cui $h_{0n}(\mathbf{x}) > \beta$, sicché $n \leq s$. Per costruzione $(0, n)$ è un segmento proprio di \mathbf{x} . Come provato, esso è contenuto in un segmento massimale di rango s che soddisfa le richieste del lemma.

Veniamo al viceversa. Il punto \mathbf{x} abbia un segmento massimale (a, b) di rango s talché $a \leq 0 < b$. Visto che $b \leq b - a < s$ non ci resta che provare che $h_{0b}(\mathbf{x}) > \beta$. Se $a = 0$ l'asserto è banale, visto che $(0, b)$ diviene un segmento proprio. Sia allora $a < 0$. Abbiamo

$$(b - a) h_{ab}(\mathbf{x}) = -a h_{a0}(\mathbf{x}) + b h_{0b}(\mathbf{x})$$

perciò

$$(c.v.d.) \quad h_{0b}(\mathbf{x}) = \frac{(b - a) h_{ab}(\mathbf{x}) + a h_{a0}(\mathbf{x})}{b} > \frac{(b - a) \beta + a \beta}{b} = \beta.$$

Dimostrazione del fatto che $m(M^*) = 0$ Consideriamo ora un qualsiasi punto \mathbf{x} dell'insieme M_s e un segmento massimale di rango s , (a, b) corrispondente a \mathbf{x} nel senso del lemma precedente. Per cui $-p \equiv a \leq 0 < b$ e $q \equiv b - a \leq s$. Abbiamo

$$\begin{aligned} 1 &\leq q \leq s \\ 0 &\leq p \leq q - 1 \end{aligned}$$

Denotiamo con δ_{pq} il segmento $(-p, -p + q)$ e con M_{pq} l'insieme dei punti di M_s che corrispondono al segmento δ_{pq} nel senso del lemma precedente. Ne viene che

$$M_s = \bigcup_{q=1}^s \bigcup_{p=0}^{q-1} M_{pq}$$

Visto che

$$h_{0q}(\mathbf{x}) = h_{-p, q-p}(U(p)\mathbf{x})$$

abbiamo che l'insieme M_{0q} evolve nell'insieme M_{pq} dopo un tempo p , perciò

$$m(M_{pq}) = m(M_{0q})$$

Gli insiemi M_{pq} con coppie p, q differenti non hanno punti in comune. Infine, se φ è una funzione di fase sommabile,

$$\int_{M_{pq}} \varphi(\mathbf{x}) dm(\mathbf{x}) = \int_{M_{0q}} \varphi(U(p)\mathbf{x}) dm(\mathbf{x})$$

Visto tutto quanto detto concludiamo

$$\begin{aligned} \int_{M_s} x_0(\mathbf{x}) dm(\mathbf{x}) &= \sum_{q=1}^s \sum_{p=0}^{q-1} \int_{M_{pq}} x_0(\mathbf{x}) dm(\mathbf{x}) = \sum_{q=1}^s \sum_{p=0}^{q-1} \int_{M_{0q}} x_0(U(p)\mathbf{x}) dm(\mathbf{x}) = \\ &= \sum_{q=1}^s \sum_{p=0}^{q-1} \int_{M_{0q}} dm(\mathbf{x}) \int_0^1 f(U(t)U(p)\mathbf{x}) dt = \\ &= \sum_{q=1}^s \sum_{p=0}^{q-1} \int_{M_{0q}} dm(\mathbf{x}) \int_p^{p+1} f(U(\tau)\mathbf{x}) d\tau = \\ &= \sum_{q=1}^s \int_{M_{0q}} dm(\mathbf{x}) \int_0^q f(U(\tau)\mathbf{x}) d\tau = \end{aligned}$$

$$\begin{aligned}
&= \sum_{q=1}^s \int_{M_{0q}} q h_{0q}(\mathbf{x}) dm(\mathbf{x}) > \beta \sum_{q=1}^s q m(M_{0q}) = \\
&= \beta \sum_{q=1}^s \sum_{p=0}^{q-1} m(M_{pq}) = \beta m(M_s)
\end{aligned}$$

Siccome la relazione vale per s grande, mandando $s \rightarrow \infty$ troviamo a secondo membro $\beta m(M^*)$ mentre a primo membro abbiamo

$$\lim_{s \rightarrow \infty} \int x_0(\mathbf{x}) \chi_{M_s} dm(\mathbf{x}) = \int x_0(\mathbf{x}) \lim_{s \rightarrow \infty} \chi_{M_s} dm(\mathbf{x}) = \int_{M^*} x_0(\mathbf{x}) dm(\mathbf{x})$$

avendo applicato il teorema di Lebesgue in ragione del fatto che $x_0(\mathbf{x})$ domina la successione di funzioni numerata da s ed è sommabile. Concludiamo

$$\int_{M^*} x_0(\mathbf{x}) dm(\mathbf{x}) \geq \beta m(M^*).$$

Lavorando in modo analogo sul limite inferiore

$$\int_{M^*} x_0(\mathbf{x}) dm(\mathbf{x}) \leq \alpha m(M^*).$$

Se $m(M^*) \neq 0$ ne ricaviamo $\beta \leq \alpha$ che è evidentemente un assurdo.

**Conclusione
della
dimostrazione
del teorema
di Birkhoff**

In definitiva il limite di h_{0n} deve esistere quasi ovunque. A questo punto, per provare il teorema di Birkhoff rimane da rimuovere il fatto che n sia intero. A questo scopo usiamo il primo lemma.

La quantità

$$\frac{1}{b} \int_0^{[b]} f(U(t)\mathbf{x}) dt$$

differisce dalla

$$\frac{1}{[b]} \int_0^{[b]} f(U(t)\mathbf{x}) dt$$

solo per un fattore tendente a 1, visto che la seconda tende a 0, anche la prima tende a 0. Poi,

$$\begin{aligned}
\left| \frac{1}{b} \int_0^b f(U(t)\mathbf{x}) dt - \frac{1}{[b]} \int_0^{[b]} f(U(t)\mathbf{x}) dt \right| &\leq \frac{1}{b} \int_{[b]}^b |f(U(t)\mathbf{x})| dt \leq \frac{1}{[b]} \int_{[b]}^{[b]+1} |f(U(t)\mathbf{x})| dt = \\
&= \frac{y_{[b]}(\mathbf{x})}{[b]} \rightarrow 0.
\end{aligned}$$

(c.v.d.) Il teorema di Birkhoff resta provato.

1.1.4 Indecomponibilità metrica

**Indipendenza
della media
temporale dal
punto iniziale
della traiettoria**

Fissiamo la seguente notazione

$$\hat{f}(\mathbf{x}) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T f(U(t)\mathbf{x}) dt$$

dove $\hat{f}(\mathbf{x})$ denota la media nel tempo dell'osservabile f sul sistema che all'istante $t = 0$ occupava lo stato \mathbf{x} . La funzione \hat{f} è ben definita su V a parte un insieme di misura nulla, perciò se assumiamo che in \mathbf{x} la \hat{f} abbia senso, ci chiediamo se sia o meno possibile mostrare che

$$\hat{f}(\mathbf{x}) = \hat{f}(U(t)\mathbf{x})$$

La cosa è del tutto ragionevole dal punto di vista fisico: ci si aspetta che la media temporale sia indipendente dallo stato iniziale e sia una prerogativa della particolare traiettoria passante per \mathbf{x} piuttosto che del punto \mathbf{x} in sé.

Dimostrazione

Proviamo l'asserto matematicamente. Fissiamo $t > 0$ e in \mathbf{x} la \hat{f} sia definita. Abbiamo allora

$$\lim_{T \rightarrow +\infty} \frac{1}{t+T} \int_0^{t+T} f(U(\tau)\mathbf{x}) d\tau = \hat{f}(\mathbf{x})$$

Inoltre,

$$\left| \frac{1}{T} \int_0^{t+T} f(U(\tau) \mathbf{x}) d\tau - \frac{1}{t+T} \int_0^{t+T} f(U(\tau) \mathbf{x}) d\tau \right| = \left| \frac{t}{T} \frac{1}{(t+T)} \int_0^{t+T} f(U(\tau) \mathbf{x}) d\tau \right| \rightarrow 0$$

per $T \rightarrow +\infty$, sicché

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^{t+T} f(U(\tau) \mathbf{x}) d\tau = \hat{f}(\mathbf{x}).$$

Tuttavia,

$$\begin{aligned} \frac{1}{T} \int_0^T f(U(\tau) U(t) \mathbf{x}) d\tau &= \frac{1}{T} \int_0^T f(U(\tau+t) \mathbf{x}) d\tau = \frac{1}{T} \int_t^{T+t} f(U(\tau) \mathbf{x}) d\tau = \\ &= \frac{1}{T} \int_0^{T+t} f(U(\tau) \mathbf{x}) d\tau - \frac{1}{T} \int_0^t f(U(\tau) \mathbf{x}) d\tau \end{aligned}$$

Nell'ultimo membro, il primo addendo tende a $\hat{f}(\mathbf{x})$, mentre il secondo tende a 0, perciò

$$\hat{f}(U(t) \mathbf{x}) \equiv \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T f(U(\tau) U(t) \mathbf{x}) d\tau = \hat{f}(\mathbf{x})$$

come volevamo. Ricapitolando

Teorema I.5
(di Birkhoff)

In un sistema hamiltoniano omogeneo governato dall'operatore di evoluzione temporale $U(t)$, la media temporale di una osservabile f , sommabile secondo Lebesgue e definita su $X \subset \Gamma$ U -invariante e di misura di Lebesgue finita, esiste per quasi ogni $\mathbf{x} \in X$. Se denotiamo tale media temporale con

$$\hat{f}(\mathbf{x}) \equiv \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T f(\mathbf{x}) dm(\mathbf{x})$$

abbiamo che \hat{f} non dipende da \mathbf{x} ma dalla traiettoria hamiltoniana passante per \mathbf{x} , cioè

$$\hat{f}(\mathbf{x}) = \hat{f}(U(t) \mathbf{x})$$

perciò \hat{f} è un integrale primo del moto.

Osservazione I.2

Risulta opportuno notare che la misura citata nel teorema di Birkhoff è quella di Lebesgue, ma nella dimostrazione i fatti richiesti circa la misura sono solamente i seguenti

- (i) la misura deve essere U -invariante;
- (ii) lo spazio X deve avere misura finita.

Se la misura è quella di Lebesgue, soddisfatta la (ii) la (i) è garantita dal teorema di Liouville. Nulla vieta, in ogni caso, di utilizzare misure diverse da quelle di Lebesgue. Per esempio se g è un integrale primo del moto sommabile, la misura $g(\mathbf{x}) dm(\mathbf{x})$ è una misura che soddisfa alle due ipotesi anche scegliendo $X \equiv \Gamma$. L'avvertenza circa la natura della misura resta valida anche per i risultati che deriveremo nel seguito.

Indecomponibilità metrica

Un caso particolarmente importante del teorema di Birkhoff si realizza quando lo spazio X è metricamente indecomponibile. Sotto questa eventualità, infatti, la media temporale dell'osservabile \hat{f} può essere sostituita dalla media spaziale.

Andiamo con ordine. Lo spazio invariante $X \subset \Gamma$ si dice **metricamente indecomponibile** se non è possibile rappresentarlo nella forma

$$X = X_1 \cup X_2$$

dove X_1 e X_2 sono insiemi **invarianti** e di misura **non nulla**.

Visto che X è un insieme invariante, esso è composto da un insieme di traiettorie complete. Se noi, in qualche modo, riusciamo a separare le traiettorie in due sottoinsiemi distinti di traiettorie complete, il fatto che X sia indecomponibile si traduce nelle due possibilità

uno dei due insiemi ha misura nulla, perciò l'altro ha misura pari a quella di X ;

nessuno dei due è misurabile.

Assumiamo allora che X sia invariante ed indecomponibile. Mostriamo allora che quasi ovunque in X ,

$$\hat{f}(\mathbf{x}) = \frac{1}{m(X)} \int_X f(\mathbf{x}') dm(\mathbf{x}')$$

La quantità a secondo membro è una media sullo spazio della funzione f e noi la denoteremo con il simbolo $\langle f \rangle$.

Il teorema che abbiamo enunciato asserisce che la media temporale $\hat{f}(\mathbf{x})$ non solo è indipendente dal dato iniziale \mathbf{x} della traiettoria, ma pure dalla traiettoria stessa, sicché la media è legata soltanto all'insieme delle traiettorie che costituiscono lo spazio indecomponibile X .

**Dimostrazione
dell'eguaglianza
delle medie
spaziali e
temporali**

Cominciamo con il dimostrare che, nelle ipotesi in cui ci siamo posti, la \hat{f} è costante quasi ovunque in X . Per assurdo, se così non fosse, esisterebbero un numero reale α e due sottoinsiemi X_1 e X_2 di X tali che

$$\hat{f}(\mathbf{x}) > \alpha, \forall \mathbf{x} \in X_1$$

$$\hat{f}(\mathbf{x}) \leq \alpha, \forall \mathbf{x} \in X_2$$

entrambi gli insiemi, essendo \hat{f} misurabile, sarebbero misurabili e di misura non nulla. Istituita questa partizione di X in due misurabili di misura non nulla, per giungere a una contraddizione, basta vedere che X_1 e X_2 sono invarianti. Ora, \hat{f} è un integrale primo del moto per quanto provato all'inizio della sottosezione, ne viene che X_1 e X_2 risultano invarianti.

In definitiva, $\hat{f}(\mathbf{x}) = a$ quasi ovunque in X . Dobbiamo adesso mostrare che $a = \langle f \rangle$. Poniamo

$$f_T(\mathbf{x}) = \frac{1}{T} \int_0^T f(U(t)\mathbf{x}) dt$$

Abbiamo

$$a = \frac{1}{m(X)} \int_X a dm(\mathbf{x}) = \frac{1}{m(X)} \int_X [a - f_T(\mathbf{x})] dm(\mathbf{x}) + \frac{1}{m(X)} \int_X f_T(\mathbf{x}) dm(\mathbf{x})$$

essendo X invariante

$$\begin{aligned} \frac{1}{m(X)} \int_X f_T(\mathbf{x}) dm(\mathbf{x}) &= \frac{1}{Tm(X)} \int_X dm(\mathbf{x}) \int_0^T f(U(t)\mathbf{x}) dt = \\ &= \frac{1}{Tm(X)} \int_0^T dt \int_X f(U(t)\mathbf{x}) dm(\mathbf{x}) = \\ &= \frac{1}{Tm(X)} \int_0^T dt \int_X f(\mathbf{x}) dm(\mathbf{x}) = \frac{1}{m(X)} \int_X f(\mathbf{x}) dm(\mathbf{x}) = \langle f \rangle \end{aligned}$$

sicché

$$a - \langle f \rangle = \frac{1}{m(X)} \int_X [a - f_T(\mathbf{x})] dm(\mathbf{x})$$

Ne abbiamo che il secondo membro **non dipende** da T e che dobbiamo dimostrare che esso è nullo. Sia $\varepsilon > 0$ fissato. Chiamiamo $X_1(T)$ l'insieme dei punti $\mathbf{x} \in X$ per cui

$$|a - f_T(\mathbf{x})| < \varepsilon$$

e $X_2(T)$ il suo complementare. Allora

$$\begin{aligned} \left| \int_X [a - f_T(\mathbf{x})] dm(\mathbf{x}) \right| &\leq \int_{X_1(T)} |a - f_T(\mathbf{x})| dm(\mathbf{x}) + \int_{X_2(T)} |a - f_T(\mathbf{x})| dm(\mathbf{x}) \leq \\ &\leq \varepsilon m(X) + |a| m(X_2(T)) + \int_{X_2(T)} |f_T(\mathbf{x})| dm(\mathbf{x}) \end{aligned}$$

Poiché $f_T(\mathbf{x})$ tende ad a quasi ovunque in X , la misura di $X_2(T)$ tende a 0 per $T \rightarrow +\infty$. Scegliamo T per cui

$$m(X_2(T)) < \varepsilon$$

Ora

$$\int_{X_2(T)} |f_T(\mathbf{x})| dm(\mathbf{x}) \leq \frac{1}{T} \int_0^T dt \int_{X_2(T)} |f(U(t)\mathbf{x})| dm(\mathbf{x}) = \frac{1}{T} \int_0^T dt \int_{U(t)X_2(T)} |f(\mathbf{x})| dm(\mathbf{x})$$

poiché

$$m(U(t)X_2(T)) = m(X_2(T)) \rightarrow 0, T \rightarrow \infty$$

si ha, applicando il teorema di Lebesgue,

$$\int_{U(t)X_2(T)} |f(\mathbf{x})| dm(\mathbf{x}) \rightarrow 0, T \rightarrow \infty$$

per cui possiamo scegliere T sufficientemente grande affinché (oltre a valere la condizione di prima)

$$\int_{U(t)X_2(T)} |f(\mathbf{x})| dm(\mathbf{x}) < \varepsilon \implies \int_{X_2(T)} |f_T(\mathbf{x})| dm(\mathbf{x}) < \varepsilon,$$

infine,

$$\left| \int_X [a - f_T(\mathbf{x})] dm(\mathbf{x}) \right| \leq \varepsilon m(X) + |a|\varepsilon + \varepsilon,$$

per l'arbitrarietà di ε , la tesi.

I.1.5 Funzioni di struttura

**Terminologia
e notazioni**

Da un punto di vista strettamente fisico, la funzione di fase più importante di un sistema è la sua energia (cioè la sua hamiltoniana), che denoteremo $H(\mathbf{x})$, dove, come sempre, $\mathbf{x} = (\mathbf{q}, \mathbf{p})$. Per un sistema isolato, la funzione $H(\mathbf{x})$ non dipende esplicitamente dal tempo ed è un integrale primo del moto. Ne viene che per ogni valore di E nell'immagine di H , la controimmagine di $\{E\}$ è un insieme invariante nello spazio delle fasi. Per semplicità e brevità ci riferiremo a questa regione come superficie a energia costante E . Considereremo poi solo i casi in cui H è limitata inferiormente in Γ , essendo questi i casi di maggiore interesse fisico. Usando l'arbitrarietà nella scelta della costante additiva con la quale si definisce l'energia potenziale, poniamo il limite inferiore di H a 0. Perciò su Γ avremo $H(\mathbf{x}) \geq 0$. Inoltre, supporremo che la regione dello spazio delle fasi per cui $H(\mathbf{x}) < E$ (E un valore positivo dell'energia) sia semplicemente connessa. Chiameremo Σ_E la superficie a energia E . Troviamo che se $E_1 < E_2$ la superficie Σ_{E_1} è tutta contenuta nella zona interna alla superficie Σ_{E_2} . Nel moto naturale del sistema ogni superficie e ogni corona tra superfici fissate è mappata in sé dall'evoluzione temporale, cioè è invariante. Per chiudere, concordiamo sul fatto che le ipotesi descritte sopra siano d'ora in poi sottintese. Come si sarà notato esse sono giustificate dalla considerazione dei sistemi fisici di maggiore interesse e familiarità.

Denotiamo con $V(E)$ il volume della parte V_E dello spazio Γ nella quale $H(\mathbf{x}) < E$. $V(E)$ è una funzione monotona che cresce da 0 a $+\infty$, mentre E descrive la medesima regione. Il volume della regione compresa tra Σ_{E_1} e Σ_{E_2} vale $V(E_2) - V(E_1)$.

**Costruzione
della misura
invariante sulla
varietà $H(\mathbf{x}) = E$**

Consideriamo un sistema isolato. Il moto si svolge nella regione invariante Σ_E , superficie a energia costante E . Per poter sfruttare il teorema di Birkhoff ci occorre definire una misura invariante sulla superficie Σ_E . Di certo sappiamo che la misura di Lebesgue è invariante in Γ per il teorema di Liouville. Cerchiamo di sfruttare questo fatto nella nostra costruzione. Innanzitutto parametrizziamo la superficie con il seguente cambiamento di coordinate

$$\mathbf{x} \mapsto \boldsymbol{\eta} \equiv (\eta_1 = H; \eta_2, \dots, \eta_{2s})$$

dove abbiamo supposto $\nabla H|_{\Sigma_E} = 0$. I punti sulla superficie $H = E$ sono individuati dalle $2s - 1$ coordinate η_2, \dots, η_{2s} . Abbiamo

$$d^{2s}\mathbf{x} = |\det J| dH d\eta_2 \dots d\eta_{2s}, \quad J \equiv \frac{\partial \mathbf{x}}{\partial \boldsymbol{\eta}}.$$

Usiamo un piccolo artificio per calcolare $\det J$. Abbiamo

$$J^{-1} \cdot (J^{-1})^t = J^{-1} \cdot (J^t)^{-1} = (J^t \cdot J)^{-1}$$

nel nostro caso,

$$(J^t \cdot J)^{-1} = J^{-1} \cdot (J^{-1})^t = \begin{pmatrix} \partial H / \partial \mathbf{x} \\ \vdots \\ \partial \eta_{2s} / \partial \mathbf{x} \end{pmatrix} (\nabla H \quad \dots \quad \nabla \eta_{2s}) =$$

$$\begin{aligned}
&= \left(\frac{|\nabla H|^2}{*} \middle| \frac{*}{(\nabla \eta_i \cdot \nabla \eta_j)_{i,j \geq 2}} \right) \\
J^t \cdot J &= \left((\partial \mathbf{x} / \partial H)^t \quad \dots \quad (\partial \mathbf{x} / \partial \eta_{2s})^t \right) \begin{pmatrix} \partial \mathbf{x} / \partial H \\ \vdots \\ \partial \mathbf{x} / \partial \eta_{2s} \end{pmatrix} = \\
&= \left(\frac{|\partial \mathbf{x} / \partial H|^2}{*} \middle| \frac{*}{(\partial \mathbf{x} / \partial \eta_i \cdot \partial \mathbf{x} / \partial \eta_j)_{i,j \geq 2}} \right)
\end{aligned}$$

Per la regola di calcolo dell'inversa

$$|\nabla H|^2 = \frac{\det (\nabla \eta_i \cdot \nabla \eta_j)_{i,j \geq 2}}{\det (J^t \cdot J)} = \frac{\det \mu}{(\det J)^2}$$

avendo posto

$$\mu_{ij} \equiv \nabla \eta_i \cdot \nabla \eta_j$$

Ne viene quindi

$$\det J = \frac{\sqrt{\det \mu}}{|\nabla H|},$$

da cui otteniamo finalmente

$$d^{2s} \mathbf{x} = \frac{dH}{|\nabla H|} \sqrt{\det \mu} d\eta_2 \dots d\eta_{2s} = \frac{dH}{|\nabla H|} d\Sigma$$

avendo notato che l'elemento di superficie $d\Sigma$ vale

$$d\Sigma = \sqrt{\det \mu} d\eta_2 \dots d\eta_{2s}.$$

Sfruttando il teorema di Liouville e il fatto che H è invariante sotto l'evoluzione temporale, concludiamo che la misura $d\Sigma / |\nabla H|$ è invariante sulla nostra varietà. La misura invariante così ottenuta, $m_{\text{mc}} = \delta(H - E) d\Sigma / |\nabla H|$ si dice **microcanonica**.

Funzione di struttura

Consideriamo adesso una funzione di fase f per cui esista l'integrale rispetto alla misura microcanonica, i.e.

$$\int_{\Sigma_E} f(\boldsymbol{\eta}) \frac{d\Sigma}{|\nabla H|}.$$

Vogliamo caratterizzare meglio la quantità detta. Abbiamo, cambiando variabile

$$\int_{V(E)} f(\mathbf{x}) d^{2s} \mathbf{x} = \int_{0 \leq H \leq E} f(H, \eta_2, \dots, \eta_{2s}) dH \frac{d\Sigma}{|\nabla H|} = \int_0^E dH \int_{\Sigma_H} f \frac{d\Sigma}{|\nabla H|}$$

Ne abbiamo che

$$\frac{d}{dE} \int_{V(E)} f(\mathbf{x}) d^{2s} \mathbf{x} \Big|_E = \int_{\Sigma_E} f \frac{d\Sigma}{|\nabla H|}$$

Definiamo adesso

$$\Omega(E) \equiv \int_{\Sigma_E} \frac{d\Sigma}{|\nabla H|}$$

misura della varietà a energia E . Ω si dice anche **funzione di struttura**. Per quanto visto sopra, posto $f = \chi_{\Sigma_E}$ ricaviamo

$$\Omega(E) = \frac{d}{dE} m(V(E))$$

Adesso rinormalizziamo la misura microcanonica ponendo

$$m_{\text{mc}}(M) = \frac{1}{\Omega(E)} \int_{\Sigma_E} f \frac{d\Sigma}{|\nabla H|}.$$

I.1.6 Sistemi composti

Sistemi separati

Consideriamo un sistema per cui la hamiltoniana H si separi in due hamiltoniane distinte H_1 e H_2 di modo che

$$H(\mathbf{x}) = H_1(\mathbf{x}_1) + H_2(\mathbf{x}_2)$$

dove $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)$ e $\mathbf{x}_i = (\mathbf{q}_i, \mathbf{p}_i)$, $i \in J_2$.

Un sistema siffatto si dice composto da due sottosistemi indipendenti, infatti le equazioni di

Composizione
delle funzioni
di struttura

Hamilton per i due raggruppamenti di variabili sono separate. La varietà a energia E si ottiene perciò come prodotto diretto delle due varietà $H_1(\mathbf{x}_1) = E_1$ e $H_2(\mathbf{x}_2) = E_2 = E - E_1$, nei due spazi $\Gamma_1 \equiv \{\mathbf{x}_1\}$ e $\Gamma_2 \equiv \{\mathbf{x}_2\}$.

Vogliamo adesso determinare la relazione che corre tra le funzioni di struttura $\Omega(E)$, $\Omega_1(E_1)$ e $\Omega_2(E_2)$. La formula che deriveremo risulterà veramente importante nella costruzione della nostra teoria.

Per far questo abbiamo bisogno di un piccolo risultato preliminare. Essendo,

$$d^{2s}\mathbf{x} = dH \frac{d\Sigma}{|\nabla H|}$$

se g è una funzione dell'energia totale, allora

$$\int_{E_1 < H < E_2} g(H(\mathbf{x})) d^{2s}\mathbf{x} = \int_{E_1}^{E_2} dH g(H) \int_{\Sigma_H} \frac{d\Sigma}{|\nabla H|} = \int_{E_1}^{E_2} dH g(H) \Omega(H).$$

Per un sistema ad hamiltoniana separata, usando il teorema di Fubini

$$\begin{aligned} m(V(E)) &= \int_{V(E)} d^{2s}\mathbf{x} = \int_{V(E)} d^{2s_1}\mathbf{x}_1 d^{2s_2}\mathbf{x}_2 = \int_{V_1(E)} d^{2s_1}\mathbf{x}_1 \int_{V_2(E-H_1)} d^{2s_2}\mathbf{x}_2 = \\ &= \int_{V_1(E)} d^{2s_1}\mathbf{x}_1 m_2(V_2(E-H_1)) = \int_0^{H_1} dH_1 \Omega_1(H_1) m_2(V_2(E-H_1)) \end{aligned}$$

dove $V_i(E)$ è l'insieme dei punti dello spazio Γ_i per i quali $E_i \leq E$, m_i è la misura di Lebesgue in Γ_i . Notato che $m_2(V_2(E-H_1))$ è nulla per $H_1 > E$, abbiamo

$$m(V(E)) = \int_0^{+\infty} dH_1 \Omega_1(H_1) m_2(V_2(E-H_1))$$

deriviamo in E ambedue i membri,

$$\Omega(E) = \int_0^{+\infty} dH_1 \Omega_1(H_1) \Omega_2(E-H_1) = (\Omega_2 * \Omega_1)(E).$$

Per induzione, nel caso di n sottosistemi si ottiene

$$\Omega(E) = \int \Omega_n \left(E - \sum_{i=1}^{n-1} H_i \right) \prod_{i=1}^{n-1} (\Omega_i(H_i) dH_i).$$

Meccanica
statistica e
sistemi separati

La meccanica statistica basa i suoi metodi sul fatto che le particelle (che possono essere riguardate come sottosistemi) si scambino tra loro l'energia posseduta tramite interazioni più o meno semplici. Dall'altra parte, si ammette le singole particelle siano descrivibili in stati di energia intensiva. In questo modo si suppone che le particelle costituiscano sottosistemi separati e che l'hamiltoniana sia data dalla somma delle hamiltoniane di singola particella. Tuttavia, così facendo, si esclude la possibilità di qualsiasi scambio energetico e si viene a una contraddizione.

La soluzione di questo paradosso si ha pensando che l'hamiltoniana è effettivamente separata ma solo al limite termodinamico (volume e numero di costituenti infiniti, densità costante). Siccome i sistemi reali sono composti da $\mathcal{N}_A \approx 6 \times 10^{23}$ particelle, ognuna occupante un volume molto piccolo rispetto al volume totale, si può assumere che il limite termodinamico sia una buona approssimazione del sistema reale, nel quale l'interazione e lo scambio energetico sono riguardati come effetti perturbativi.

Su tutti questi aspetti che risultano fondanti sia della meccanica statistica, sia della termodinamica, ritorneremo diffusamente in seguito.

I.2 Il problema ergodico

La fondazione della meccanica statistica non può prescindere da una discussione del problema ergodico. Questo problema è ben lungi dall'essere risolto, cionondimeno la sua focalizzazione consente di porre sotto una luce critica aspetti della meccanica quantistica che sono solitamente assunti senza una previa discussione.

I.2.1 Interpretazione delle quantità fisiche in meccanica statistica

I valori delle quantità fisiche che possiamo misurare sul sistema sono univocamente determinate

dallo stato del sistema stesso, e questo è descritto da un punto nello spazio delle fasi. Ne viene che le osservabili sono funzioni definite sullo spazio delle fasi. Se f è la funzione su Γ che rappresenta la quantità che vogliamo misurare, si pone il problema di come mettere in relazione la funzione al risultato di una misura. Siccome una misura avviene in un certo tempo, e siccome è ragionevole assumere che i tempi di misura siano lunghi rispetto alla scala dei tempi dell'evoluzione microscopica, è ragionevole assumere che il risultato della misura sia ben approssimato dalla media temporale della f ,

$$\bar{f} = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T f(U(t)\mathbf{x}) dt$$

Il teorema di Birkhoff ci assicura che \bar{f} è ben definito, ma, a meno che non assumiamo indecomponibilità metrica, non ci consente di determinare in modo semplice il valore di \bar{f} . Va da sé che per noi è anzi impossibile determinare \bar{f} dalla definizione, visto che appare complicato determinare il dato iniziale \mathbf{x} che comprende $6N_A$ componenti e poi risolvere le equazioni di Hamilton.

L'unica cosa che solitamente sappiamo è l'energia del sistema e perciò che l'orbita del sistema viene a trovarsi su una varietà $2s - 1$ dimensionale nota. Se tale superficie fosse metricamente indecomponibile, le nostre difficoltà sarebbero di colpo superate, poiché accadrebbe che

$$\bar{f} = \int f(\mathbf{x}) d\mu(\mathbf{x}) = \langle f \rangle$$

dove μ è una misura invariante sulla varietà a energia E .

Nei testi di meccanica statistica, solitamente, si introduce la media spaziale come interpretazione della misura delle osservabili in modo del tutto ingiustificato. In assenza di indecomponibilità metrica non esiste una giustificazione di questa interpretazione, se non a posteriori: poiché i risultati che così si ottengono sono molto buoni, allora l'impostazione scelta è corretta, anche se non la sola possibile. In altri termini, associare la media spaziale alla previsione del risultato di una misura è una decisione arbitraria, plausibile e che porta a risultati buoni.

In quanto segue, ci occuperemo di chiarire perché la decisione detta è *plausibile*. Evidentemente l'unico modo per fare questo è di indagare al meglio delle nostre possibilità il senso che ha porre $\bar{f} = \langle f \rangle$. Di certo la sostituzione $\bar{f} = \langle f \rangle$ è, non solo desiderabile, ma inevitabile se vogliamo fare qualche conto...

Una volta assunta l'eguaglianza di media spaziale e media temporale, ci si porrà il problema di come effettuare il calcolo delle medie spaziali, ma questo aspetto sarà enormemente semplificato dall'uso di metodi statistici: ce ne occuperemo nei prossimi capitoli.

1.2.2 Integrali del moto

Il problema ergodico

Il problema della giustificazione teorica della sostituzione della media temporale con la media spaziale, si dice **problema ergodico**. Quasi sempre si considerano medie spaziali prese sulla varietà a energia costantee uno dei punti importanti è quello di capire perché questo avviene.

La varietà a energia costante

Solitamente, il ragionamento che si fa è il seguente: siccome l'energia è un integrale primo del moto, ogni traiettoria giace sulla superficie a energia costante Σ_E ; i valori delle funzioni di fase in punti che non appartengono a Σ_E sono del tutto ininfluenti, perciò non devono essere considerati, affinché sia possibile che \bar{f} e $\langle f \rangle$ siano eguali (almeno approssimativamente), di qui la media spaziale effettuata sulla sola Σ_E .

Questo ragionamento sarebbe corretto se non contenesse un evidente elemento di arbitrarità: perché preferire l'energia a tutti gli altri integrali del moto (localmente ne esistono $2s - 1$)?

Perché scegliere la varietà a energia costante

Ora, si potrebbe dare subito una risposta molto pratica al quesito: solitamente (lo abbiamo detto prima) l'unico integrale noto è l'energia, perciò la domanda fatta non ha molto senso. Tuttavia, questo è vero in pratica, ma non in linea di principio, inoltre è possibile immaginare sistemi in cui si conservi, ad esempio, una componente del momento angolare, per cui rispondere alla domanda di sopra, non è solo un gioco inutile.

Se non specifichiamo Σ_E ma andiamo ad eseguire la media su Γ , sorge subito il problema che non tutte le osservabili ammettono integrale finito, a meno che non si definisca un peso

(integrale del moto, affinché la misura che ne deriva sia invariante), ma l'introduzione di un peso è tanto arbitraria, quanto la specifica di eseguire l'integrazione sulla varietà a energia E . In ogni caso, fissato un peso opportuno (senza preoccuparci troppo dell'arbitrarietà della sua scelta), possiamo determinare la media dell'energia, $\langle H \rangle$, e chiederci quale senso fisico assegnare a tale quantità. Evidentemente è impossibile pensare che $\langle H \rangle$ possa essere vicina al valor medio dell'energia per ogni possibile preparazione del sistema. Quello che impariamo è che il solo tentativo di attribuire all'energia di un sistema un qualsiasi valore fissato, come $\langle H \rangle$, in un qualsiasi modo, contraddice la realtà. Ne viene che, almeno per l'energia, la previa riduzione di Γ a una varietà almeno contenuta in Σ_E è inevitabile.

Vediamo adesso perché non dobbiamo prestare attenzione a tutti gli altri integrali del moto. Un'analisi accurata della situazione reca i seguenti

- (i) come vedremo poi, la maggior parte delle osservabili fisiche con le quali si ha a che fare in meccanica statistica hanno una struttura ben precisa che rende i valori di queste funzioni variare molto poco su Σ_E , perciò la loro media temporale risulterà bene approssimata dalla media spaziale.
- (ii) Sia I un qualsiasi integrale del moto indipendente dall'energia. Considerando I come una funzione di fase, per quanto detto al punto precedente, ha senso sostituirne la media temporale con quella su Σ_E . Comunque, se I non soddisfa le condizioni dette, esso non ha, di norma, alcuna interpretazione fisica e perciò non ci interessa.
- (iii) A volte è possibile però che esistano quantità fisicamente rilevanti e che non soddisfano le condizioni di cui al primo punto: in questi casi è veramente il caso di ridurre la varietà Σ_E a una varietà ridotta in cui, oltre a H , anche I sia costante.

Integrali controllabili

Se chiamiamo **controllabili** gli integrali di cui al terzo punto, abbiamo che, in presenza di k integrali controllabili (integrali il cui valore, come quello dell'energia, è facilmente controllabile nella preparazione del sistema) la regione da considerare nello spazio Γ è una varietà $2s - k$ dimensionale. Un esempio dell'effetto che hanno gli integrali controllabili sulle statistiche è dato dallo spin che impone in un caso o l'altro l'uso della Bose-Einstein o della Fermi-Dirac. Il fatto che nella maggior parte dei casi, l'unico integrale controllabile sia l'energia, fa sì che solitamente si vadano a considerare correttamente le medie sulla superficie Σ_E .

Qualunque sia la varietà sulla quale si riconosce di dover lavorare, il valore degli altri integrali aventi una precisa interpretazione fisica (e perciò un qualche interesse) è approssimativamente fissato sulla varietà, perciò la posizione $\bar{f} = \langle f \rangle$ è plausibile. Sarà nostro proposito mostrare la validità dei tre punti descritti per chiudere la questione in modo definitivo.

I.2.3 Indecomponibilità metrica delle varietà ridotte

Problema ergodico e indecomponibilità metrica

Il teorema di Birkhoff stabilisce l'esistenza delle medie temporali e ne stabilisce l'eguaglianza con le medie spaziali nel caso in cui si riscontri indecomponibilità metrica dell'insieme $X \subset \Gamma$ che stiamo considerando. Si vede facilmente che vale anche il viceversa. Se assumiamo cioè che per quasi ogni traiettoria (dato iniziale) media temporale e media spaziale e contemporaneamente assumiamo che esistano X_1 e X_2 disgiunti, invarianti, misurabili, di misura non nulla, e aventi unione pari a X , abbiamo che la funzione χ_{X_1} non può avere la stessa media temporale della media spaziale.

Ne deriva che l'indecomponibilità metrica è equivalente all'ergodicità. In questo modo abbiamo trasferito il problema ergodico al problema di stabilire se o meno una varietà a energia costante sia metricamente indecomponibile.

Integrali del moto e indecomponibilità metrica

Sia φ un integrale del moto indipendente dall'energia, se il valore di φ fosse del tutto stabilito su ciascuna varietà a energia costante (di modo che, di certo, media spaziale e temporale coinciderebbero), allora φ sarebbe una funzione dell'energia, contrariamente all'assunzione fatta. Di conseguenza φ non può rimanere costante su ogni Σ_E e nemmeno può farlo a meno di un insieme di misura nulla, se φ è continua.

Prendiamo una Σ_E sulla quale φ non sia costante, allora possiamo trovare un numero reale α ,

per cui le controimmagini di $\varphi > \alpha$ e $\varphi \leq \alpha$ abbiano misura non nulla. Essendo φ un integrale del moto, abbiamo così che Σ_E è metricamente decomponibile e medie temporali e spaziali non possono coincidere.

Questo semplice argomento sembra mostrare che qualunque ipotesi ergodica (ipotesi volta a stabilire che media temporale e spaziale coincidano) è destinata ad essere fallimentare e con ciò che il problema ergodico è risolto in modo negativo.

Necessità di riformulare il problema ergodico

In effetti, nelle condizioni che abbiamo imposto, questo è vero, tuttavia, basta indebolire un po' le richieste per tornare a dare un senso (e la speranza di una soluzione) al problema ergodico. Una possibile variazione all'argomento precedente, peraltro suggerita da quanto visto nella sottosezione precedente, è quella di non considerare ogni possibile integrale del moto φ nel ragionamento di sopra, ma di restringerci solo a una classe di integrali aventi diretta interpretazione fisica e, dunque, un interesse (si tratta di restringere l'algebra delle osservabili, alle sole osservabili che hanno un senso).

Stati di un sistema e osservabili come funzioni di fase normali

Precisiamo quanto detto. Per un fisico è facile pensare a un sistema il cui stato non possa essere ragionevolmente descritto da un singolo punto nello spazio delle fasi, ma da un insieme di punti dello spazio del tutto equivalenti per l'individuazione dello stato fisico stesso. D'altra parte, è ovvio che una qualsiasi osservabile che abbia senso fisico preciso (che sia **veramente** un'osservabile!) deve assumere gli stessi valori su insiemi di punti equivalenti perché descrittivi del medesimo stato. Tali funzioni (vere osservabili, abbiamo detto) si dicono **funzioni normali**. Poiché il problema ergodico ha senso solo per le osservabili effettive (per le altre funzioni non ha alcun senso confrontare medie spaziali e medie temporali), dobbiamo restringerci a considerare solo funzioni di fase normali. Ne viene che la condizione di indecomponibilità metrica cessa di essere necessaria e viene chiaramente sostituita da una condizione più ampia.

Chiameremo **normale** ogni partizione di un dato insieme X (superficie a energia costante, per esempio), ogni coppia (X_1, X_2) di insiemi misurabili, disgiunti, invarianti e di misura non nulla tali che insieme di punti equivalenti sono tutti contenuti o in X_1 o in X_2 . Un insieme X per cui non sia possibile effettuare una suddivisione normale, si dice **metricamente indecomponibile in senso esteso**.

In questo modo è facile riottenere (ripercorrendo le dimostrazioni fatte nel caso più forte) il seguente

Teorema I.6

Condizione necessaria e sufficiente affinché ogni funzione normale sommabile assuma media temporale su quasi ogni traiettoria eguale alla media spaziale su X è che X sia metricamente indecomponibile in senso esteso.

Conclusione

A questo punto, si ottiene che Σ_E è metricamente indecomponibile se non esistono integrali primi indipendenti dall'energia che siano normali. Se questo avviene si deve restringere Σ_E alla varietà ridotta in cui tutti gli integrali primi normali sono fissati.

1.2.4 Formulazione senza uso di indecomponibilità metrica

Lo studio del problema ergodico di Birkhoff è troppo generale per i nostri scopi, dal momento che esso rimane valido per sistemi formati da un qualsivoglia numero di particelle. Noi, invece, siamo interessati principalmente ai sistemi contenenti un gran numero di gradi di libertà, qualche \mathcal{N}_A , perciò adesso vogliamo specificare questa richiesta.

Come più volte detto, noi vogliamo giustificare l'eguaglianza di media spaziale e media temporale soltanto per quelle funzioni di fase che hanno un preciso senso fisico. Per queste, come abbiamo detto, si riscontra un comportamento molto particolare: a prescindere dalla indecomponibilità metrica, sono quasi costanti sulle varietà a energia fissata, di modo che le medie spaziali e temporali debbono ritenersi, in qualche senso, vicine.

La ragione di questo particolare comportamento è da ricercarsi in parte nel fatto che i sistemi statistici contengono un gran numero di gradi di libertà, in parte nelle caratteristiche di queste osservabili. Di norma, assumeremo come osservabili di nostro interesse le *funzioni somma*, date cioè dalla somma di funzioni di singola particella.

Detto questo, vediamo in che modo è possibile dire che funzioni quasi costanti (costanti a parte un insieme di misura finita) abbiano media temporale vicina alla media spaziale. Assumiamo che $f(\mathbf{x})$ sulla superficie Σ_E sia una quantità molto vicina a A . Allora, è ragionevole supporre che, se la funzione non differisce troppo da A nell'insieme di piccola misura in cui è diversa da A stesso, la quantità

$$I = \frac{1}{\Omega(E)} \int_{\Sigma_E} |f(\mathbf{x}) - A| \frac{d\Sigma}{|\nabla H|}$$

è piccola. Assumiamo per semplicità $A = 0$. Poniamo poi

$$f_T = \frac{1}{T} \int_0^T f(U(t)\mathbf{x}) dt, \quad \bar{f}(\mathbf{x}) = \lim_{T \rightarrow \infty} f_T(\mathbf{x})$$

e chiamiamo M_α l'insieme dei punti di Σ_E per i quali $|\bar{f}(\mathbf{x})| > \alpha$, inoltre, chiamiamo M_α^T l'insieme dei punti per cui $|f_T(\mathbf{x})| > \alpha/2$.

Poiché \bar{f} è il limite di f_T , abbiamo che per T sufficientemente ampio

$$m_{\text{mc}}(M_\alpha^T) > \frac{1}{2} m_{\text{mc}}(M_\alpha).$$

Per provare questa disuguaglianza consideriamo $(M_\alpha^T)^c$ complementare di M_α^T . Se $\mathbf{x} \in M_\alpha$ e $\mathbf{x} \in (M_\alpha^T)^c$, abbiamo

$$|f_T(\mathbf{x}) - \bar{f}(\mathbf{x})| > \frac{\alpha}{2}$$

Conseguentemente, abbiamo

$$m(M_\alpha \cap (M_\alpha^T)^c) \rightarrow 0$$

Perciò, per T grande

$$m(M_\alpha \cap M_\alpha^T) > \frac{1}{2} m(M_\alpha)$$

da cui

$$m(M_\alpha^T) > \frac{1}{2} m(M_\alpha).$$

Allora

$$\begin{aligned} \frac{\alpha}{4} m_{\text{mc}}(M_\alpha) &< \frac{1}{\Omega(E)} \int_{M_\alpha^T} |f_T(\mathbf{x})| \frac{d\Sigma}{|\nabla H|} \leq \frac{1}{T\Omega(E)} \int_0^T dt \int_{M_\alpha^T} |f(U(t)\mathbf{x})| \frac{d\Sigma}{|\nabla H|} = \\ &= \frac{1}{T\Omega(E)} \int_0^T dt \int_{U(t)M_\alpha^T} |f(U(t)\mathbf{x})| \frac{d\Sigma}{|\nabla H|} \leq \frac{1}{T\Omega(E)} \int_0^T dt \int_{\Sigma_E} |f(\mathbf{x})| \frac{d\Sigma}{|\nabla H|} = I \end{aligned}$$

da cui segue

$$m(M_\alpha) < \frac{4I}{\alpha}$$

e preso $\alpha = \sqrt{I}$ si ottiene

$$m(M_{\sqrt{I}}) < 4\sqrt{I}$$

per cui la misura dell'insieme sul quale \bar{f} è più grande di \sqrt{I} rispetto alla media spaziale è inferiore a $4\sqrt{I}$.

Vedremo in seguito che la dispersione \tilde{I} delle osservabili principali scala come N , perciò, essendo, per la disuguaglianza di Schwarz,

$$I \leq \tilde{I}^{1/2}$$

abbiamo che I scala come $N^{1/2}$, perciò scelto $\alpha = I^{3/2}$, l'insieme dei punti per cui

$$|\bar{f}(\mathbf{x}) - A| > KN^{3/4}$$

ha una misura che scala come $I^{-1/2} = O(N^{-1/4})$.

Questa rozza disussione mette in luce come, almeno nei casi fondamentali, sia possibile fare a meno della teoria ergodica per ritenere almeno approssimativamente, che le medie spaziali ben rappresentino quelle temporali.

Teoria delle probabilità e meccanica statistica

In questo capitolo presentiamo la riduzione del problema della meccanica statistica alla teoria delle probabilità, assumendo appunto che la media spaziale sia una buona interpretazione (rigorosa nel caso ergodico, approssimata altrimenti) del risultato di una misura fisica sul sistema.

L'uso del teorema del limite centrale ci porterà alla legge di Boltzmann e all'introduzione dell'insieme canonico di Gibbs. Nei capitoli successivi ci occuperemo dei gas ideali e della fondazione della termodinamica.

II.1 Riduzione del problema alla teoria delle probabilità

II.1.1 Legge fondamentale di distribuzione

Coordinate
canoniche
come variabili
aleatorie

Consideriamo il vettore \mathbf{x} come una variabile aleatoria. Assumiamo che l'energia E del sistema sia fissata una volta per tutte, sicché i possibili valori di \mathbf{x} sono situati sulla varietà Σ_E . La probabilità che \mathbf{x} vada a cadere nell'insieme $M \subset \Sigma_E$ sia

$$p_{\text{mc}}(M) = \frac{1}{\Omega(E)} \int_M \frac{d\Sigma}{|\nabla H|}$$

dove $\Omega(E)$ è la funzione di struttura di Σ_E .

Legge di
distribuzione
fondamentale

La legge di distribuzione per la variabile \mathbf{x} postulata si chiama **legge di distribuzione fondamentale del sistema**. Sia adesso φ una funzione di fase misurabile. La probabilità che si verifichi la diseuguaglianza $\varphi < a$ varrà

$$p_{\text{mc}}(\varphi < a) = \frac{1}{\Omega(E)} \int_{\varphi < a} \frac{d\Sigma}{|\nabla H|}$$

perciò ogni funzione misurabile può essere considerata una quantità aleatoria con una ben precisa legge di distribuzione e il cui valore di aspettazione matematico vale

$$E(\varphi) = \frac{1}{\Omega(E)} \int_{\Sigma_E} \varphi \frac{d\Sigma}{|\nabla H|}$$

Tale valore di aspettazione coincide con quella che abbiamo chiamato, in precedenza, media spaziale della funzione φ e che nel caso ergodico (diciamo pure in senso esteso) eguaglia la media temporale e con ciò il risultato di una misura sperimentale di φ del sistema.

Interpretazione
fisica della
riduzione alla
teoria delle
probabilità

Se φ è la funzione caratteristica di $M \subset \Sigma_E$, allora il suo valore di aspettazione è pari a $p_{\text{mc}}(M)$. Se φ è una funzione ergodica, tale probabilità coincide con il tempo medio relativo speso dal punto \mathbf{x} nell'insieme M durante tutto il moto del punto \mathbf{x} .

La legge di distribuzione fondamentale ci consente di introdurre una terminologia e una metodologia probabilistica nella risoluzione del problema del calcolo delle medie spaziali.

II.1.2 Legge di distribuzione per una componente del sistema

Sistemi separati

Il sistema G ammetta suddivisione in due componenti G_1 , descritta dal gruppo di variabili canoniche \mathbf{x}_1 , e G_2 , descritta da \mathbf{x}_2 . La legge di distribuzione per G , cioè per la variabile \mathbf{x} , determina univocamente la legge di distribuzione per le componenti G_1 e G_2 (i.e., per le \mathbf{x}_1 e \mathbf{x}_2).

Legge di
distribuzione
per le
componenti
singole

Sia M_1 un insieme misurabile in Γ_1 (spazio delle \mathbf{x}_1). Sia M un insieme di tutti i punti di

Γ per i quali \mathbf{x}_1 appartiene a M_1 per ogni $\mathbf{x} \in M$ (\mathbf{x}_1 essendo la proiezione sulle prime $2s_1$ coordinate, di modo che M_1 è il proiettato, sulle prime $2s_1$ coordinate, di M).

Ora,

$$p_{\text{mc}}(\mathbf{x}_1 \in M_1) = p_{\text{mc}}(\mathbf{x} \in M) = \frac{1}{\Omega(E)} \int_{M \cap \Sigma_E} \frac{d\Sigma}{|\nabla H|} = \frac{1}{\Omega(E)} \int_{\Sigma_E} \chi_M \frac{d\Sigma}{|\nabla H|}$$

ne viene, per quanto visto al capitolo precedente, che

$$p_{\text{mc}}(\mathbf{x}_1 \in M_1) = \frac{1}{\Omega(E)} \int_{\Sigma_E} \chi_M \frac{d\Sigma}{|\nabla H|} = \frac{1}{\Omega(E)} \frac{d}{dE} \int_{V(E)} \chi_M d^{2s} \mathbf{x}$$

Poiché χ_M non dipende dalla \mathbf{x}_2 , abbiamo

$$\int_{V(E)} \chi_M d^{2s} \mathbf{x} = \int_{V_1(E)} \chi_M d^{2s_1} \mathbf{x}_1 \int_{V_2(E-H_1)} d^{2s_2} \mathbf{x}_2$$

Poiché se $H_1 > E$ il secondo integrale è nullo, possiamo sostituire il dominio d'integrazione del primo integrale con tutto lo spazio Γ_1 ottenendo

$$\int_{V(E)} \chi_M d^{2s} \mathbf{x} = \int_{\Gamma_1} \chi_M m_2(V_2(E-H_1)) d^{2s_1} \mathbf{x}_1 = \int_{M_1} m_2(V_2(E-H_1)) d^{2s_1} \mathbf{x}_1$$

di conseguenza

$$\frac{d}{dE} \int_{V(E)} \chi_M d^{2s} \mathbf{x} = \int_{M_1} \Omega_2(E-H_1) d^{2s_1} \mathbf{x}_1$$

Infine,

$$p_{\text{mc}}(\mathbf{x}_1 \in M_1) = \frac{1}{\Omega(E)} \int_{M_1} \Omega_2(E-H_1(\mathbf{x}_1)) d^{2s_1} \mathbf{x}_1 = \int_{M_1} \frac{\Omega_2(E-H_1(\mathbf{x}_1))}{\Omega(E)} d^{2s_1} \mathbf{x}_1$$

La formula ricavata mostra come nella componente G_1 sia definita una densità di probabilità pari a $\Omega_2(E-H_1)/\Omega(E)$, dove Ω_2 è la funzione di struttura di G_2 .

In particolare se φ è una funzione di fase dipendente da \mathbf{x}_1 soltanto, il suo valore di aspettazione (media spaziale) viene calcolato integrando sull'intero Γ_1 :

$$\langle \varphi \rangle = E(\varphi) = \frac{1}{\Omega(E)} \int_{\Gamma_1} \varphi(\mathbf{x}_1) \Omega_2(E-H_1(\mathbf{x}_1)) d^{2s_1} \mathbf{x}_1.$$

Infine, sostituendo a φ l'osservabile H_1 , abbiamo

$$\langle H_1 \rangle = \frac{1}{\Omega(E)} \int_{\Gamma_1} H_1(\mathbf{x}_1) \Omega_2(E-H_1(\mathbf{x}_1)) d^{2s_1} \mathbf{x}_1$$

Legge di distribuzione per l'energia di una singola componente

Adesso occupiamoci della determinazione della funzione di distribuzione per la variabile H_1 stessa. Abbiamo visto che la variabile aleatoria \mathbf{x}_1 ammette come densità di probabilità

$$\frac{\Omega_2(E-H_1(\mathbf{x}_1))}{\Omega(E)},$$

perciò la probabilità che risulti $E' < H_1 < E''$ vale

$$p(E' < H_1 < E'') = \frac{1}{\Omega(E)} \int_{E' < H_1 < E''} \Omega_2(E-H_1(\mathbf{x}_1)) d^{2s_1} \mathbf{x}_1$$

cambiando variabile come abbiamo imparato nel primo capitolo,

$$\begin{aligned} p(E' < H_1 < E'') &= \frac{1}{\Omega(E)} \int_{E'}^{E''} dH_1 \Omega_2(E-H_1) \int_{\Sigma_{H_1}} \frac{d\Sigma_1}{|\nabla H_1|} = \\ &= \frac{1}{\Omega(E)} \int_{E'}^{E''} dH_1 \Omega_2(E-H_1) \Omega_1(H_1) \end{aligned}$$

di modo che la quantità H_1 è soggetta alla seguente densità di probabilità

$$\frac{\Omega_2(E-H_1) \Omega_1(H_1)}{\Omega(E)}. \quad (\text{II.1})$$

In particolare,

$$\langle H_1 \rangle = \frac{1}{\Omega(E)} \int_0^{+\infty} dH_1 H_1 \Omega_2(E-H_1) \Omega_1(H_1)$$

dove abbiamo ammesso che l'integrazione si protraesse fino a $+\infty$, poiché, comunque, per $H_1 > E$, $\Omega_2 = 0$.

Siccome nelle applicazioni ci capiterà spesso di avere a che fare con sistemi separati e funzioni dipendenti solo da gruppi di variabili, e siccome ci occorreranno, come abbiamo appena visto, le funzioni di struttura dell'insieme complessivo e dei sottosistemi, elaboreremo, nel seguito, un metodo per approssimare convenientemente tali funzioni di struttura.

II.1.3 Funzioni generatrici

Consideriamo un sistema G la cui funzione di struttura Ω sia soggetta alle solite condizioni: sia positiva e monotonamente crescente per $E > 0$ in modo illimitato. Inoltre, sia pari a zero per $E \leq 0$ e sia continua. Di certo è possibile applicare alla Ω la trasformata di Laplace per ottenere

$$\Phi(\alpha) = \int e^{-\alpha E} \Omega(E) dE$$

che assumiamo avere ascissa di convergenza nulla. La funzione $\Phi(\alpha)$ si dice **funzione generatrice** della funzione di struttura $\Omega(E)$.

In questa sottosezione passiamo in rassegna le principali proprietà della Φ :

Proposizione II.1 $\Phi(\alpha)$ è una funzione positiva, monotona decrescente, infinitesima all'infinito e infinita all'origine.

Dimostrazione La positività è un fatto del tutto ovvio, come anche la decrescenza: $e^{-\alpha E}$ è decrescente in α . Il limite per $\alpha \rightarrow +\infty$ si trova nel modo seguente:

$$e^{-\bar{\alpha} E} \Omega(E) \in L^1$$

per $\bar{\alpha} > 0$. Ora, per ogni $\alpha > \bar{\alpha}$ risulta, per ogni E

$$e^{-\alpha E} \Omega(E) < e^{-\bar{\alpha} E} \Omega(E) \in L^1$$

perciò, per il teorema di Lebesgue

$$\lim_{\alpha \rightarrow +\infty} \Phi(\alpha) = \int \lim_{\alpha \rightarrow +\infty} e^{-\alpha E} \Omega(E) dE = 0.$$

Per l'ultimo asserto, abbiamo

$$\begin{aligned} \Phi(\alpha) &= \int_0^{+\infty} e^{-\alpha E} \Omega(E) dE \geq \int_{\bar{E}}^{+\infty} e^{-\alpha E} \Omega(E) dE \geq \Omega(\bar{E}) \int_{\bar{E}}^{+\infty} e^{-\alpha E} dE = \\ &= \Omega(\bar{E}) \frac{e^{-\alpha \bar{E}}}{\alpha} \rightarrow +\infty, \alpha \rightarrow 0. \end{aligned}$$

(c.v.d.)

Come sappiamo la trasformata di Laplace di una funzione Laplace-trasformabile (come abbiamo ammesso che sia Ω , per $\alpha > 0$) è olomorfa nel semipiano di convergenza, perciò Φ è infinitamente derivabile per $\alpha > 0$ e si ha

Proposizione II.2 La funzione generatrice $\Phi(\alpha)$ estesa al piano complesso è olomorfa nel semipiano reale, perciò è infinitamente derivabile per $\alpha > 0$ e risulta

$$\Phi^{(n)}(\alpha) = (-1)^n \int E^n e^{-\alpha E} \Omega(E) dE.$$

Proposizione II.3 La derivata seconda logaritmica di $\Phi(\alpha)$ è positiva per $\alpha > 0$.

Dimostrazione Basta fare il conto,

$$\begin{aligned} \frac{d^2}{d\alpha^2} \log(\Phi(\alpha)) &= \frac{d}{d\alpha} \frac{\Phi'}{\Phi} = \frac{\Phi'' \Phi - (\Phi')^2}{\Phi^2} = \\ &= \frac{1}{\Phi(\alpha)} \int \left(E + \frac{\Phi'(\alpha)}{\Phi(\alpha)} \right)^2 e^{-\alpha E} \Omega(E) dE > 0 \end{aligned}$$

Infatti, l'ultimo membro è pari a

$$\frac{1}{\Phi(\alpha)} \int E^2 e^{-\alpha E} \Omega(E) dE + \frac{2 \Phi'}{\Phi} \int E e^{-\alpha E} \Omega(E) dE + \frac{(\Phi')^2}{\Phi^3} \Phi$$

$$(c.v.d.) \quad = \frac{\Phi''}{\Phi} - 2 \frac{(\Phi'(\alpha))^2}{\Phi^2} + \frac{(\Phi'(\alpha))^2}{\Phi^2}.$$

Proposizione II.4 *L'equazione*

$$-\frac{\Phi'(\alpha)}{\Phi(\alpha)} = a$$

ha una singola soluzione per ogni $a > 0$.

Dimostrazione Definiamo la funzione

$$\Phi_a(\alpha) \equiv e^{a\alpha} \Phi(\alpha)$$

Essa tende all'infinito per $\alpha \rightarrow 0$, in forza della prima proposizione. Anche la funzione $\log \Phi_a$ possiede la medesima proprietà. Quest'ultima funzione è pure convessa avendo derivata seconda positiva:

$$\frac{d^2}{d\alpha^2} \log \Phi_a(\alpha) = \frac{d^2}{d\alpha^2} (a\alpha + \log \Phi(\alpha)) = \frac{d^2}{d\alpha^2} \log \Phi(\alpha) > 0.$$

Poiché, infine, la funzione $\log \Phi_a$ diviene infinita per $\alpha \rightarrow \infty$, oltre che, come detto, per $\alpha \rightarrow 0$, se ne conclude che essa deve ammettere uno e un solo minimo nell'intervallo $(0, +\infty)$. Se α è l'ascissa di tale minimo concludiamo

$$(c.v.d.) \quad 0 = \frac{d}{d\alpha} \log \Phi_a(\alpha) = a + \frac{\Phi'(\alpha)}{\Phi(\alpha)}.$$

Infine, sempre facendo appello alle proprietà delle trasformate di Laplace, abbiamo, poiché in un sistema composto $\Omega = \Omega_2 * \Omega_1$, per le corrispondenti funzioni generatrici

$$\Phi = \Phi_2 \Phi_1$$

Proposizione II.5 *La funzione generatrice Φ di un sistema composto da due sottosistemi aventi rispettivamente funzione generatrice Φ_1 e Φ_2 vale*

$$\Phi(\alpha) = \Phi_2(\alpha) \Phi_1(\alpha).$$

Per esempio se G è un gas consistente di N molecole identiche aventi ciascuna $\phi(\alpha)$ come funzione generatrice, allora la funzione generatrice dell'intero sistema risulta

$$\Phi(\alpha) = \phi^N(\alpha).$$

Un'ultima osservazione, nel capitolo precedente abbiamo dimostrato che

$$\int_{\Gamma} f(H(\mathbf{x})) d^{2s}\mathbf{x} = \int f(H) \Omega(H) dH$$

perciò, posto $f(H) \equiv e^{-\alpha H}$, otteniamo

$$\Phi(\alpha) = \int e^{-\alpha H} \Omega(H) dH = \int_{\Gamma} e^{-\alpha H(\mathbf{x})} d^{2s}\mathbf{x}.$$

II.1.4 Leggi di distribuzione coniugate

**Famiglia della
distribuzioni
coniugate**

Poniamo la seguente definizione, preso $\alpha > 0$,

$$U^{(\alpha)}(E) \equiv \begin{cases} \frac{1}{\Phi(\alpha)} e^{-\alpha E} \Omega(E), & E \geq 0 \\ 0, & E < 0 \end{cases}$$

Siccome risulta in modo ovvio

$$U^{(\alpha)}(E) \geq 0, \quad \int U^{(\alpha)}(E) dE = 1$$

la quantità $U^{(\alpha)}(E)$ può essere riguardata come una densità di probabilità per l'energia, per ogni $\alpha > 0$. La famiglia di funzioni di densità che otteniamo per il sistema G al variare di $\alpha \in \mathbb{R}^+$ si dice famiglia delle **leggi di distribuzione coniugate** del sistema G .

Viceversa, la funzione di struttura $\Omega(E)$ può essere ottenuta a partire da un qualsiasi elemento della famiglia $U^{(\alpha)}$ essendo

$$\Omega(E) = \Phi(\alpha) e^{\alpha E} U^{(\alpha)}(E).$$

Valore di aspettazione per le distribuzioni coniugate

Il valor medio della distribuzione $U^{(\alpha)}(E)$ vale

$$a \equiv \int E U^{(\alpha)}(E) dE = \frac{1}{\Phi(\alpha)} \int E e^{-\alpha E} \Omega(E) dE = -\frac{\Phi'(\alpha)}{\Phi(\alpha)} = -\frac{d}{d\alpha} \log \Phi(\alpha)$$

siccome, assegnato $a > 0$ esiste una e una sola radice dell'equazione

$$a = -\frac{d}{d\alpha} \log \Phi(\alpha)$$

si ricava il seguente

Teorema II.6 Per ogni numero positivo a , si trova sempre una e una sola distribuzione $U^{(\alpha)}$ appartenente alla famiglia delle distribuzioni coniugate che abbia a come valore di aspettazione.

Per quanto concerne la dispersione di una quantità E distribuita secondo la $U^{(\alpha)}$ abbiamo

$$\begin{aligned} b &\equiv \int E^2 U^{(\alpha)}(E) dE - a^2 = \frac{1}{\Phi(\alpha)} \int E^2 e^{-\alpha E} \Omega(E) dE - a^2 = \\ &= \frac{\Phi''(\alpha) \Phi(\alpha) - (\Phi'(\alpha))^2}{\Phi^2(\alpha)} = \frac{d^2}{d\alpha^2} \log \Phi(\alpha). \end{aligned}$$

Composizione delle distribuzioni coniugate

Ricordiamo che la legge di composizione per le funzioni di struttura è

$$\Omega(E) = (\Omega_2 * \Omega_1)(E)$$

per cui

$$\begin{aligned} \Phi(\alpha) e^{\alpha E} U^{(\alpha)}(E) &= \Phi_1(\alpha) \Phi_2(\alpha) \int dE_1 e^{\alpha(E-E_1)} U_2^{(\alpha)}(E-E_1) e^{\alpha E_1} U_1^{(\alpha)}(E_1) \\ \Phi(\alpha) e^{\alpha E} U^{(\alpha)}(E) &= \Phi(\alpha) e^{\alpha E} (U_2^{(\alpha)} * U_1^{(\alpha)})(E) \end{aligned}$$

sicché si conclude

$$U^{(\alpha)}(E) = (U_2^{(\alpha)} * U_1^{(\alpha)})(E).$$

Per induzione si trova

$$U^{(\alpha)}(E) = \int \prod_{k=1}^{n-1} (U_k^{(\alpha)}(E_k) dE_k) U_n^{(\alpha)}\left(E - \sum_{k=1}^{n-1} E_k\right).$$

La formula trovata è notevolissima. Infatti consideriamo una sequenza di variabili aleatorie $\{x_k\}$ **positive** e **indipendenti** distribuite secondo le leggi u_k , allora la distribuzione u della somma delle x_k è data proprio dalla

$$u_k\left(x = \sum_{k=1}^n x_k\right) = \int \prod_{k=1}^{n-1} (u_k(x_k) dx_k) u_n\left(x - \sum_{k=1}^{n-1} x_k\right)$$

perciò la distribuzione coniugata di un dato sistema può essere derivata dalle corrispondenti distribuzioni delle sue n componenti allo stesso modo in cui si determina la distribuzione della somma di n quantità aleatorie indipendenti a partire dalle distribuzioni dei termini individuali.

Sistemi composti da un grande numero di componenti

Nel nostro caso x è l'energia totale (come si vede si assume che l'energia sia additiva, il che è vero nel limite termodinamico), x_k è l'energia della singola componente, u_k è $U_k^{(\alpha)}$ e u è $U^{(\alpha)}$. Siccome il teorema del limite centrale dà una stima analitica di u , noi dobbiamo utilizzarlo per trovare U^α . Poiché la stima è tanto migliore quanto più è grande il numero di sottosistemi, la determinazione di $U^{(\alpha)}$ sarà ottima nel limite termodinamico. Come vedremo, non dovremo fare assunzioni troppo forti sulla natura fisica delle componenti, ma il nostro risultato dipenderà in modo determinante solo dal fatto che l'energia è additiva e che il numero di componenti è molto grande. Ne viene, lo ribadiamo, che la nostra teoria dovrà funzionare al limite termodinamico che è, sperimentalmente una ottima approssimazione per i sistemi reali.

II.2 Applicazioni del teorema del limite centrale

II.2.1 Il teorema del limite centrale

Come probabilmente noto al lettore, il teorema del limite centrale stabilisce un'espressione approssimata per la legge di distribuzione che governa la somma di un grande numero di quantità aleatorie mutuamente indipendenti.

Per il momento ci limitiamo ad enunciare il teorema in una forma non proprio generale, ma sufficiente per i nostri scopi (rimandandone la dimostrazione).

Teorema II.7
(del limite centrale)

Consideriamo una sequenza di quantità casuali indipendenti $\{u_k(x)\}_{k \geq 1}$ aventi funzioni caratteristiche $\{g_k(t)\}_{k \geq 1}$, dove

$$g_k(t) \equiv \int e^{itx} u_k(x) dx, \quad k \geq 1$$

Posto che esistano le seguenti quantità

$$a_k \equiv \int x u_k(x) dx;$$

$$b_k \equiv \int (x - a_k)^2 u_k(x) dx;$$

$$c_k \equiv \int |x - a_k|^3 u_k(x) dx;$$

$$d_k \equiv \int (x - a_k)^4 u_k(x) dx;$$

$$e_k \equiv \int |x - a_k|^5 u_k(x) dx;$$

e assunto che ciascuna distribuzione soddisfi le seguenti condizioni

(i) per ogni $k \geq 1$, u_k sia differenziabile ed esista L talché

$$\int |u'_k(x)| dx < L;$$

(ii) esistano due costanti positive $\alpha < \beta$ per cui, per ogni $k \geq 1$, si abbia

$$\alpha < b_k < \beta, \quad c_k < \beta, \quad d_k < \beta, \quad e_k < \beta;$$

(iii) esistano due costanti positive λ e τ tali che per $|t| \leq \tau$ valga, per ogni $k \geq 1$,

$$|g_k(t)| > \lambda;$$

(iv) per ogni intervallo (c_1, c_2) nella semiretta reale positiva esista un numero $\rho = \rho(c_1, c_2) < 1$ per cui, per ogni $t \in (c_1, c_2)$ sia

$$|g_k(t)| < \rho, \quad \forall k \geq 1;$$

allora, posto

$$A_n \equiv \sum_{k=1}^n a_k; \quad B_n \equiv \sum_{k=1}^n b_k$$

e denominata $U_n(x)$ la densità di probabilità per la somma delle prime n $u_k(x)$, abbiamo, per $n \rightarrow \infty$

$$U_n(x) = \frac{1}{\sqrt{2\pi B_n}} \exp \left[-\frac{(x - A_n)^2}{2B_n} \right] + \varepsilon_n(x)$$

dove

$$\varepsilon_n(x) = \begin{cases} o \left(\frac{1 + |x - A_n|}{n^{3/2}} \right), & |x - A_n| < 2 \log^2 n \\ o(1/n), & \forall x \end{cases}$$

**Teorema del
limite centrale
e meccanica
statistica**

Noi useremo il teorema del limite centrale per determinare la forma di $U^{(\alpha)}(E)$ di un dato sistema G sotto l'unica assunzione che tale sistema consista di un grande numero di sottosistemi

g_1, \dots, g_N . Il sistema g_k abbia funzione di struttura $\omega_k(E_k)$, funzione generatrice $\phi_k(E_k)$, e funzione coniugata $u_k^{(\alpha)}(E_k)$. Abbiamo già mostrato nella sezione precedente che per noi le u_k del teorema del limite centrale sono le $u_k^{(\alpha)}(E_k)$. Rimane evidentemente l'arbitrarietà di α che utilizzeremo in seguito per semplificare le formule.

Realizzazione delle ipotesi del teorema in meccanica statistica

Detto come intendiamo usare il teorema del limite centrale per i nostri scopi, mostriamo che le ipotesi del teorema sono plausibili con le condizioni che abbiamo sulle nostre funzioni $u_k^{(\alpha)}$. Il punto è che le richieste sulle u_k nel teorema del limite centrale hanno tutte il carattere dell'uniformità (esistenza di certe costanti che abbiano lo stesso comportamento per ogni $k \geq 1$, cioè **uniformemente**). Ora, nel dominio della meccanica statistica (il limite termodinamico), i sottosistemi g_k (atomi, molecole...) sono tutti identici oppure si presentano un piccolo numero di grandi raggruppamenti di componenti identiche (tipicamente nelle miscele). Di conseguenza le funzioni di struttura e conseguentemente le distribuzioni coniugate di queste componenti formano un insieme entro al quale tutti gli elementi sono identici oppure esiste un piccolo numero di sottoinsiemi formati tutti da elementi identici tra loro. Questo significa che, in queste condizioni, tutte le ipotesi di uniformità del teorema sono soddisfatte se sono soddisfatte dalla distribuzione coniugata di un singolo sottosistema. Preoccupiamoci di questo. Usualmente le funzioni di struttura ω_k sono analitiche e moderatamente crescenti, perciò, poiché

$$u_k^{(\alpha)}(x) = \frac{1}{\phi_k(\alpha)} e^{-\alpha x} \omega_k(x)$$

la condizione (i) è sempre soddisfatta. Ne viene subito anche l'esistenza di momenti finiti di ogni ordine, viste le osservazioni di sopra sull'uniformità, si conclude che anche l'ipotesi (ii) è sempre soddisfatta. La terza condizione richiede che g_k non si annulli per t troppo piccoli il che accade di certo se g_k non si annulla proprio in 0 (abbiamo già discusso il discorso sull'uniformità) visto che g_k è continua, ma questo accadrebbe solo se u_k avesse integrale nullo, il che è assurdo. Per quanto concerne la quarta richiesta abbiamo

$$|g_k(t)| = \left| \int e^{itx} u_k(x) dx \right| \leq \int u_k(x) dx = 1, \quad \forall t \in \mathbb{R}$$

ed essendo $u_k(x)$ continua, se $t > 0$, vale la disuguaglianza stretta, cioè

$$|g_k(t)| < 1$$

si conclude che per qualsiasi $c_1, c_2 > 0$

$$\rho \equiv \max_{(c_1, c_2)} |g_k(t)| < 1$$

dove ρ è ben definito grazie alla continuità di g_k .

Come si vede il teorema del limite centrale è veramente adatto a risolvere il nostro problema!

Applicazione del teorema del limite centrale

Non ci resta che applicare il teorema del limite centrale. Se consideriamo il nostro sistema come composto da N sottosistemi (molecole, ad esempio), poniamo, per quanto visto nella sezione precedente,

$$A_N \equiv -\frac{d}{d\alpha} \log \Phi(\alpha), \quad B_N \equiv \frac{d^2}{d\alpha^2} \log \Phi(\alpha)$$

ed otteniamo quindi

$$\Omega(H) = \Phi(\alpha) e^{\alpha H} \left[\frac{1}{\sqrt{2\pi \frac{d^2}{d\alpha^2} \log \Phi(\alpha)}} \exp \left[-\frac{(H + d/d\alpha \log \Phi(\alpha))^2}{2 \frac{d^2}{d\alpha^2} \log \Phi(\alpha)} \right] + \varepsilon_N(H) \right]$$

Per semplificare le formule, se E è l'energia totale del nostro sistema, poniamo $\alpha \equiv \theta$ dove θ è l'unica radice dell'equazione

$$-\frac{d}{d\alpha} \log \Phi(\alpha) = E$$

allora

$$A_N = E, \quad B_N = \frac{d^2}{d\alpha^2} \log \Phi(\alpha) \Big|_{\alpha=\theta} \equiv B$$

e otteniamo

$$\Omega(H) = \Phi(\theta) e^{\theta H} \left[\frac{1}{\sqrt{2\pi B}} \exp \left[-\frac{(H-E)^2}{2B} \right] + \varepsilon_N(H) \right] \quad (\text{II.2})$$

In particolare, per $H = E$, abbiamo

$$\Omega(E) = \Phi(\theta) e^{\theta E} \left[\frac{1}{\sqrt{2\pi B}} + o(n^{-3/2}) \right] \quad (\text{II.3})$$

II.2.2 La legge di Boltzmann

Piccola nota

In questo capitolo ricaveremo la ben nota legge di Boltzmann, ci preme sottolineare che la derivazione della stessa avviene dopo 32 pagine di dimostrazioni, a differenza di quanto avviene in altri testi, in cui questa legge viene giustificata all'antiquato modo di Boltzmann, oppure soltanto postulata: il tutto molto in breve!

Sistemi composti

Consideriamo un sistema G composto da due sottosistemi G_1 e G_2 . Come abbiamo avuto modo di dimostrare nella precedente sezione, nello spazio delle fasi Γ_1 si ha una legge di distribuzione data dalla densità di probabilità

$$\frac{\Omega_2(E - H_1(\mathbf{x}_1))}{\Omega(E)}$$

dove Ω_2 è la funzione di struttura di G_2 e dove E è l'energia totale del sistema G .

Divisione in un sistema molto piccolo e uno molto grande

Assumiamo, come abbiamo fatto più volte, che il sistema G sia formato da N sottosistemi identici molto piccoli rispetto a G e che chiameremo brevemente **molecole** (si immagini pure che G sia un gas formato da N molecole). Il sottosistema G_1 sia composto da N_1 molecole, laddove G_2 da N_2 , con

$$N_1 + N_2 = N.$$

Supponiamo, ora, che G_1 sia molto più piccolo di G , addirittura possiamo pensare che $N_1 = 1$ (N sia molto grande, come abbiamo già detto più volte). È molto facile rendersi conto che, se

$$A_i = - \left. \frac{d}{d\alpha} \log \Phi_i(\alpha) \right|_{\alpha=\theta}, \quad B_i = \left. \frac{d^2}{d\alpha^2} \log \Phi_i(\alpha) \right|_{\alpha=\theta}, \quad i \in J_2,$$

abbiamo

$$A_1 + A_2 = E; \quad B_1 + B_2 = B$$

infatti,

$$E = - \left. \frac{d}{d\alpha} \log (\Phi_2(\alpha) \Phi_1(\alpha)) \right|_{\alpha=\theta} = - \left. \frac{d}{d\alpha} \log (\Phi_2(\alpha)) \right|_{\alpha=\theta} - \left. \frac{d}{d\alpha} \log (\Phi_1(\alpha)) \right|_{\alpha=\theta} = A_1 + A_2$$

e analogamente per B . Similmente, se con a_k denotiamo

$$a_k = - \left. \frac{d}{d\alpha} \log (\phi_k(\alpha)) \right|_{\alpha=\theta}$$

abbiamo che A_1, A_2 e A (come $B_{1,2}, B$) sono dati dalle somme degli a_k (o dei b_k) delle molecole che li compongono. Ne viene che tutte queste quantità scalano come il numero di molecole, $A_i = O(N_i)$, etc.

Distribuzione per la parte piccola del sistema

Se G_1 è dato da un numero molto basso di molecole, ad esempio $N_1 = 1$, abbiamo

$$N_1 = o(N), \quad N \rightarrow \infty$$

e così

$$A_1 = o(E), \quad B_1 = o(B), \quad N \rightarrow \infty$$

Adesso,

$$\Omega(E) = \frac{\Phi(\theta) e^{\theta E}}{(2\pi B)^{1/2}} [1 + o(1)]$$

essendo

$$B^{1/2} o\left(\frac{1}{n^{3/2}}\right) = n^{1/2} o\left(\frac{1}{n^{3/2}}\right) = n^{3/2} o\left(\frac{1}{n^{3/2}}\right) \frac{1}{n} = o(1)$$

Tenendo conto del fatto che $E - H_1 - A_2 = A_1 - H_1$ e che $N_2 \simeq N$ si ha

$$\Omega_2(E - H_1) = \Phi_2(\theta) e^{\theta(E-H_1)} \left[\frac{\exp \left[-(H_1 - A_1)^2 / 2B_2 \right]}{\sqrt{2\pi B}} + o\left(\frac{1}{N}\right) \right]$$

dove $A_1 = o(E)$. Se ci limitiamo a considerare valori di H_1 per cui $H_1 - A_1 = o(N^{1/2})$, troviamo

$$\Omega_2(E - H_1) = \Phi_2(\theta) e^{\theta(E-H_1)} \left[\frac{1+o(1)}{\sqrt{2\pi B}} + o\left(\frac{1}{N}\right) \right] = \frac{\Phi_2(\theta) e^{\theta(E-H_1)}}{\sqrt{2\pi B}} [1+o(1)]$$

Infine, la distribuzione che otteniamo nello spazio Γ_1 vale

$$\frac{\Omega_2(E - H_1)}{\Omega(E)} = \frac{e^{-\theta H_1}}{\Phi_1(\theta)} [1+o(1)] \quad (\text{II.4})$$

valida per $H_1 - A_1 = o(N^{1/2})$.

Legge di Boltzmann

La formula (II.4) è proprio la legge di distribuzione di Boltzmann valida nello spazio delle fasi della singola molecola.

Interpretazione fisica della distribuzione coniugata

In accordo alla (II.1), la distribuzione di H_1 è data da

$$\frac{\Omega_1(H_1)\Omega_2(E - H_1)}{\Omega(E)}$$

ma preso ancora $H_1 - A_1 = o(N^{1/2})$ abbiamo che la densità di probabilità per H_1 risulta

$$\frac{\Omega_1(H_1) e^{-\theta H_1}}{\Phi_1(\theta)} [1+o(1)].$$

Va notato che abbiamo ottenuto, per la distribuzione approssimata dell'energia di una componente piccola del sistema, l'espressione esatta per la sua distribuzione coniugata

$$U_1^{(\theta)}(H_1) = \frac{\Omega_1(H_1) e^{-\theta H_1}}{\Phi_1(\theta)}$$

Affinchè questo accada, è naturalmente fondamentale che sia scelto il valore θ per α . Ad ogni modo, vediamo che la distribuzione coniugata per una piccola componente, ammette una semplice interpretazione fisica, presa per $\alpha = \theta$ essa è la distribuzione in energia della componente piccola.

Distribuzione in energia per la singola molecola

Quando G_1 è una singola molecola, allora $A_1 = a_1$ rimane costante al variare di N , e la formula (II.4) si applica uniformemente per H_1 compreso tra limiti costanti arbitrari (l'ampiezza dell'intervallo attorno ad A_1 , che è fisso, in cui vale la formula, $|H_1 - A_1| = o(N^{1/2})$, può essere fatta crescere quanto si vuole aumentando N).

La probabilità che una molecola abbia un'energia tra E' ed E'' vale

$$\int_{E'}^{E''} \frac{\omega_i(x) e^{-\theta x}}{\phi_i(x)} dx [1+o(1)]$$

perciò il numero di molecole che ci aspettiamo con energia tra E' e E'' vale

$$\sum_{i=1}^N \int_{E'}^{E''} \frac{\omega_i(x) e^{-\theta x}}{\phi_i(x)} dx + o(N)$$

II.2.3 Valori medi delle funzioni additive

Distribuzione nello spazio di singola molecola

Sia G_1 formato da una sola molecola. Dal momento che la numerazione delle molecole è irrilevante, sia $G_1 \equiv g_1$, perciò siano ω_1 e ϕ_1 , rispettivamente, la funzione di struttura e la funzione generatrice per G_1 .

Adesso operiamo un cambiamento di notazione: sia $\mathbf{X} \in \mathbb{R}^{2s}$ il vettore delle variabili dinamiche dell'intero sistema (anziché \mathbf{x}) e denotiamo con \mathbf{x}_i le variabili dinamiche della molecola i -esima. Una funzione di fase $f(\mathbf{X})$ che dipenda in realtà soltanto da \mathbf{x}_1 può essere interpretata come funzione di fase nello spazio delle fasi della singola molecola g_i . Denoteremo con γ_i lo spazio delle fasi $\{\mathbf{x}_i\}$ di singola molecola.

In γ_1 è definita la densità di probabilità

$$\frac{\Omega_2(E - H_1(\mathbf{x}_1))}{\Omega(E)}$$

dove adesso Ω_2 è la funzione di struttura per $G \setminus g_1$ (cioè per l'intero sistema privato di una molecola) e $H_1(\mathbf{x}_1)$ è la hamiltoniana relativa alla sola molecola considerata. Presa $f(\mathbf{x}_1)$ il

suo valor medio vale

$$\langle f \rangle = \int_{\gamma_1} f(\mathbf{x}_1) \frac{\Omega_2(E - H_1(\mathbf{x}_1))}{\Omega(E)} d^{2s_1} \mathbf{x}_1$$

dove s_1 indica il numero di gradi di libertà per la singola molecola. Ammettiamo che la funzione f sia un o -piccolo di H_1^k per $H_1 \rightarrow \infty$.

Valor medio per una funzione nello spazio γ_1 di singola molecola

Vogliamo usare le formule approssimate dal teorema del limite centrale per calcolare approssimativamente il valor medio di f . Dividiamo lo spazio γ_1 in due parti: γ'_1 in cui $H_1(\mathbf{x}_1) < \log^2(N)$ e $\gamma''_1 = (\gamma'_1)^c$. Poniamo

$$I' \equiv \int_{\gamma'_1} f(\mathbf{x}_1) \frac{\Omega_2(E - H_1(\mathbf{x}_1))}{\Omega(E)} d^{2s_1} \mathbf{x}_1$$

$$I'' \equiv \int_{\gamma''_1} f(\mathbf{x}_1) \frac{\Omega_2(E - H_1(\mathbf{x}_1))}{\Omega(E)} d^{2s_1} \mathbf{x}_1$$

sicché $\langle f \rangle = I' + I''$. Fissato, come al solito,

$$a_1 \equiv - \frac{d}{d\alpha} \log \phi_1 \Big|_{\alpha=\theta}$$

$$b_1 \equiv \frac{d^2}{d\alpha^2} \log \phi_1 \Big|_{\alpha=\theta}$$

abbiamo

$$\begin{aligned} \Omega_2(E - H_1) &= \Phi_2(\theta) e^{\theta(E - H_1)} \left[\frac{\exp\left[-(H_1 - a_1)^2 / 2(B - b_1)\right]}{\sqrt{2\pi(B - b_1)}} + o\left(\frac{1}{N}\right) \right] < \\ &< \Phi_2(\theta) e^{\theta(E - H_1)} \left[\frac{1}{\sqrt{2\pi(B - b_1)}} + o\left(\frac{1}{N}\right) \right] \end{aligned}$$

da cui,

$$\Omega_2(E - H_1) < \frac{\Phi_2(\theta) e^{\theta(E - H_1)}}{\sqrt{2\pi B}} \left[\sqrt{\frac{B}{B - b_1}} + o(1) \right]$$

per $N \rightarrow \infty$ la quantità entro parentesi tende a 1, perciò per N sufficientemente grande sarà inferiore a 2,

$$\Omega_2(E - H_1) < 2 \frac{\Phi_2(\theta) e^{\theta(E - H_1)}}{\sqrt{2\pi B}}$$

D'altra parte, come abbiamo visto,

$$\Omega(E) = \frac{\Phi(\theta) e^{\theta E}}{\sqrt{2\pi B}} [1 + o(1)]$$

perciò, per N sufficientemente grande,

$$\Omega(E) > \frac{1}{2} \frac{\Phi(\theta) e^{\theta E}}{\sqrt{2\pi B}}$$

Infine, per N sufficientemente grande

$$\frac{\Omega_2(E - H_1)}{\Omega(E)} < 4 \frac{e^{-\theta H_1}}{\phi_1(\theta)}$$

conseguentemente,

$$|I''| \leq \int_{\gamma''_1} |f(\mathbf{x}_1)| \frac{\Omega_2(E - H_1)}{\Omega(E)} d^{2s_1} \mathbf{x}_1 < C \int_{\gamma''_1} \frac{H_1^k e^{-\theta H_1}}{\phi_1(\theta)} d^{2s_1} \mathbf{x}_1$$

(come si vede la supposizione che f non cresca più rapidamente di una certa potenza dell'energia consente di concludere che l'integrale esiste). Cambiando variabile,

$$\begin{aligned} |I''| &< \frac{C}{\phi_1(\theta)} \int_{\log^2 N}^{+\infty} H_1^k e^{-\theta H_1} \omega_1(H_1) dH_1 = \frac{C}{\phi_1(\theta)} \int_{\log^2 N}^{+\infty} e^{-\theta H_1/2} H_1^k e^{-\theta H_1/2} \omega_1(H_1) dH_1 < \\ &< \frac{C}{\phi_1(\theta)} \exp\left(-\frac{\theta}{2} \log^2 N\right) \int_{\log^2 N}^{+\infty} H_1^k e^{-\theta H_1/2} \omega_1(H_1) dH_1 \end{aligned}$$

Poiché l'integrale tende a 0 per $N \rightarrow \infty$, per N sufficientemente grande,

$$|I''| < \exp\left(-\frac{\theta}{2} \log^2 N\right) < \frac{1}{N}$$

Pensiamo adesso a I' . Abbiamo, per $H_1 < a_1 + \log^2 N$

$$\begin{aligned} \Omega_2(E - H_1) &= \Phi_2(\theta) e^{\theta(E - H_1)} \left[1 + o\left(\frac{1 + (H_1 - a_1)^2}{N}\right) \right] \frac{1}{\sqrt{2\pi B}} \\ \Omega(E) &= \Phi_2(\theta) e^{\theta E} \left[1 + o\left(\frac{1}{N}\right) \right] \frac{1}{\sqrt{2\pi B}} \end{aligned}$$

sicché

$$\frac{\Omega_2(E - H_1)}{\Omega(E)} = \frac{e^{-\theta H_1}}{\phi_1(\theta)} \left[1 + o\left(\frac{1 + (H_1 - a_1)^2}{N}\right) \right]$$

e, in definitiva,

$$\begin{aligned} I' &= \int_{\gamma'_1} f(\mathbf{x}_1) \frac{e^{-\theta H_1}}{\phi_1(\theta)} \left[1 + o\left(\frac{1 + (H_1 - a_1)^2}{N}\right) \right] d^{2s_1} \mathbf{x}_1 = \\ &= \int_{\gamma'_1} f(\mathbf{x}_1) \frac{e^{-\theta H_1}}{\phi_1(\theta)} d^{2s_1} \mathbf{x}_1 + o\left(\frac{1}{N}\right) \end{aligned}$$

Usando il risultato ottenuto per I'' ,

$$\langle f \rangle = \int_{\gamma_1} f(\mathbf{x}_1) \frac{e^{-\theta H_1}}{\phi_1(\theta)} d^{2s_1} \mathbf{x}_1 + o\left(\frac{1}{N}\right)$$

Valor medio per una funzione dell'energia

In particolare, per $f = f(H_1)$, abbiamo

$$\langle f \rangle = \int f(H_1) \frac{e^{-\theta H_1}}{\phi_1(\theta)} \omega_1(H_1) dH_1 + o\left(\frac{1}{N}\right) = \int f(H_1) u_1^{(\theta)}(H_1) dH_1 + o\left(\frac{1}{N}\right)$$

che enfatizza il ruolo della distribuzione coniugata nel caso di singola molecola.

Dunque,

$$\begin{aligned} \langle H_1 \rangle &= \int H_1 u_1^{(\theta)}(H_1) dH_1 + o\left(\frac{1}{N}\right) = a_1 + o\left(\frac{1}{N}\right) \\ \langle (H_1 - a_1)^2 \rangle &= \int (H_1 - a_1)^2 u_1^{(\theta)}(H_1) dH_1 + o\left(\frac{1}{N}\right) = b_1 + o\left(\frac{1}{N}\right) \end{aligned}$$

A questo punto siamo in grado di determinare i valori medi per ciò che riguarda le funzioni additive, infatti il valor medio di una somma è **sempre** pari alla somma dei valori medi. Queste ultime sono del tipo

$$f(\mathbf{X}) = \sum_{i=1}^N f_i(\mathbf{x}_i)$$

e perciò si ottiene

$$\langle f \rangle = \sum_{i=1}^N \langle f_i \rangle = \sum_{i=1}^N \int_{\gamma_i} f(\mathbf{x}_i) \frac{e^{-\theta H_i}}{\phi_i(\theta)} d^{2s_i} \mathbf{x}_i + o(1).$$

Esempio II.1 Vogliamo calcolare il numero medio di molecole con energia compresa tra due limiti fissati, $0 \leq \alpha < \beta$. Tale quantità è data dal valore medio della funzione di fase

$$f(\mathbf{X}) = \sum_{i=1}^N f_i(\mathbf{x}_i)$$

dove

$$f_i(\mathbf{x}_i) = \chi_{\{\alpha < H_i(\mathbf{x}_i) < \beta\}}$$

perciò

$$\langle N(\alpha, \beta) \rangle = \langle f \rangle = \sum_{i=1}^N \int_{\alpha}^{\beta} \frac{e^{-\theta H_i}}{\phi_i(\theta)} \omega_i(H_i) dH_i + o(1) = \sum_{i=1}^N \int_{\alpha}^{\beta} u_i^{(\theta)}(H_i) dH_i + o(1)$$

In particolare, se le molecole risultano tutte eguali,

$$\frac{\langle N(\alpha, \beta) \rangle}{N} = \int_{\alpha}^{\beta} u^{(\theta)}(H) dH + o\left(\frac{1}{N}\right).$$

■

Esempio II.2 Calcoliamo il valor medio dell'energia posseduta da una componente molto grande del sistema. Se N_1 è il numero di molecole di tale componente e poniamo

$$f_i(\mathbf{x}_i) = \begin{cases} H_i(\mathbf{x}_i), & i \leq N_1 \\ 0, & \text{altrimenti} \end{cases}$$

L'energia di tale componente vale

$$E_1 = f(\mathbf{X}) = \sum_{i=1}^N f_i(\mathbf{x}_i),$$

cioè

$$\langle E_1 \rangle = \sum_{i=1}^{N_1} a_1 + o(1)$$

Non possiamo qui utilizzare lo stesso metodo per ottenere la dispersione, perché le energie di molecole differenti in G_1 non sono indipendenti e in questo caso la somma delle dispersioni non è eguale alla somma delle dispersioni (ci ritorneremo in seguito). .

■

II.2.4 Legge di distribuzione per l'energia di una componente grande

Distribuzione dell'energia per un sottosistema grande

In questa sottosezione vogliamo determinare la legge di distribuzione per l'energia di una componente grande del sistema. Denotiamo questa componente con G_1 . Sia G_2 la sua complementare. Le notazioni siano quelle usuali, in particolare l'indice 1 si riferisce a G_1 e l'indice 2, a G_2 .

Abbiamo

$$A_1 = \sum_{k=1}^{N_1} a_k, \quad A_2 = E - A_1$$

$$B_1 = \sum_{k=1}^{N_1} b_k, \quad B_2 = B - B_1.$$

La densità di probabilità per H_1 vale

$$\frac{\Omega_1(H_1)\Omega_2(E-H_1)}{\Omega(E)},$$

ma, come abbiamo più volte visto,

$$\Omega_1(H_1) = \Phi_1(\theta) e^{\theta H_1} \left[\frac{\exp\left[-(H_1 - A_1)^2/2B_1\right]}{\sqrt{2\pi B_1}} + o\left(\frac{1}{N}\right) \right]$$

$$\Omega_2(E - H_1) = \Phi_2(\theta) e^{\theta(E-H_1)} \left[\frac{\exp\left[-(H_1 - A_1)^2/2B_2\right]}{\sqrt{2\pi B_2}} + o\left(\frac{1}{N}\right) \right]$$

$$\Omega(E) = \Phi(\theta) e^{\theta E} \left[\frac{1}{\sqrt{2\pi B}} + o\left(\frac{1}{N^{3/2}}\right) \right]$$

Posto,

$$B^* \equiv \frac{B_1 B_2}{B}$$

troviamo

$$\frac{\Omega_1(H_1)\Omega_2(E-H_1)}{\Omega(E)} = \left[\frac{1}{\sqrt{4\pi^2 B_1 B_2}} \exp\left[-\frac{(H_1 - A_1)^2}{2B_1} - \frac{(H_1 - A_1)^2}{2B_2}\right] + o(1/N) \right] \frac{\sqrt{2\pi B}}{1 + o(1)} =$$

$$\begin{aligned}
 &= \sqrt{\frac{1}{2\pi} \frac{B}{B_1 B_2}} \exp \left[-\frac{(H_1 - A_1)^2}{2B^*} \right] + o\left(\frac{1}{N}\right) = \\
 &= \frac{1}{\sqrt{2\pi B^*}} \exp \left[-\frac{(H_1 - A_1)^2}{2B^*} \right] + o\left(\frac{1}{N}\right)
 \end{aligned}$$

Nota sulla correlazione delle energie delle molecole

Ne deriva che al limite per $N \rightarrow \infty$, la distribuzione dell'energia per una componente grande è data da una gaussiana centrata in A_1 e con dispersione B^* . Notiamo che

$$B^* = \frac{B_1 B_2}{B_1 + B_2} \neq B_1 = \sum_{k=1}^{N_1} b_k$$

cioè che la dispersione della somma non è la somma delle dispersioni. D'altronde, le energie delle singole molecole non possono essere indipendenti, poiché, in ogni caso, la loro somma deve essere pari a E , energia totale.

II.2.5 Gas monoatomico ideale

Descrizione del modello di gas monoatomico ideale

Come esempio di applicazione dei metodi appresi finora, consideriamo il più semplice esempio di insieme meccanico statistico: il gas monoatomico ideale. Il sistema G sia formato da N molecole g_1, \dots, g_N semplicemente rappresentabili come punti materiali (perciò che differiscono al più nella massa). Come sempre l'energia totale del sistema sarà la somma delle singole energie delle molecole di modo che l'energia d'interazione delle molecole stesse sia posta eguale a zero. Come appare chiaro questa è un'approssimazione brutale che non rispecchia la realtà, a parte nei casi di grande diluizione etc., eppure è solo in questi termini che possiamo applicare tutto quanto visto. Assumeremo che il nostro gas sia racchiuso in un recipiente di volume V costituito da pareti perfettamente elastiche. il modo formale in cui rendiamo conto della presenza delle pareti è nell'aggiunta di un campo esterno alla hamiltoniana di singola molecola, che diventa

$$H_i(\mathbf{q}_i, \mathbf{p}_i) = \frac{|\mathbf{p}_i|^2}{2m_i} + U_i(\mathbf{q}_i)$$

dove U_i è eguale in forma a U_j per ogni coppia di indici. Se l'interazione tra le pareti e le molecole avviene solo quando la distanza mutua è molto piccola e se supponiamo che le pareti siano del tutto impenetrabili, per qualsiasi velocità delle particelle incidenti, possiamo schematizzare le pareti stesse come barriere infinitamente alte e, a scopo essere semplici, supporremo che U_i sia costante entro V e infinito sulla superficie della scatola. In realtà, dovendo essere U continua avremmo dovuto assumere che, in uno spazio molto ristretto, U_i salisse dal valore costante interno all'infinito. Tuttavia, la barriera discontinua è un'approssimazione buona e tale da semplificare notevolmente i conti, perciò la accettiamo senz'altro. Per concludere, poniamo (arbitrariamente, ma in accordo alle convenzioni da noi precedentemente stabilite) $U_i(\mathbf{q}_i) = 0$ per \mathbf{q}_i rappresentante un punto interno alla scatola.

Visto che fuori dalla scatola l'energia potenziale è infinita, lo spazio delle fasi γ_i accessibile ad ogni molecola è dato dal prodotto diretto del volume interno alla scatola con lo spazio degli impulsi.

Calcolo esatto della funzione di struttura

La misura dello spazio in Γ (prodotto diretto dei γ_i) a energia $H = \sum H_i$ inferiore a E è data da

$$V(E) = \int_{H < E} d^{2s} \mathbf{X} = V^N \int_{\sum_i |\mathbf{p}_i|^2 / 2m_i < E} \prod_{i=1}^N d^3 \mathbf{p}_i$$

Cambiamo coordinate e poniamo

$$\mathbf{u}_i = \frac{\mathbf{p}_i}{\sqrt{2m_i}}$$

allora

$$\int_{\sum_i |\mathbf{p}_i|^2 / 2m_i < E} \prod_{i=1}^N d^3 \mathbf{p}_i (2m_i)^{3/2} = \int_{\sum_i |\mathbf{u}_i|^2 < E} \prod_{i=1}^N d^3 \mathbf{u}_i = \left(2^N \prod_{i=1}^N m_i \right)^{3/2} \frac{\pi^{3N/2}}{\Gamma((3N/2) + 1)} E^{3N/2}$$

In definitiva,

$$V(E) = V^N \left(\prod_{i=1}^N m_i \right)^{3/2} \frac{(2\pi)^{3N/2}}{\Gamma((3N/2) + 1)} E^{3N/2}$$

sicché la funzione di struttura risulta **esattamente**

$$\Omega(E) = \frac{d}{dE} V(E) = \frac{3}{2} N V^N \left(\prod_{i=1}^N m_i \right)^{3/2} \frac{(2\pi)^{3N/2}}{\Gamma((3N/2) + 1)} E^{(3N/2)-1}$$

Come abbiamo visto, nel caso del gas ideale, non abbiamo bisogno degli sviluppi asintotici per determinare la funzione di struttura. Comunque, per illustrare l'uso delle formule approssimate, ricaviamo la $\Omega(E)$ così come fornita dal teorema del limite centrale e la confrontiamo con quella esatta ottenuta sopra.

**Formula
asintotica per
la funzione
di struttura**

Cominciamo con il costruire la funzione generatrice,

$$\begin{aligned} \Phi(\alpha) &= \int \Omega(x) e^{-\alpha x} dx = \frac{3}{2} N V^N \left(\prod_{i=1}^N m_i \right)^{3/2} \frac{(2\pi)^{3N/2}}{\Gamma((3N/2) + 1)} \int x^{(3N/2)-1} e^{-\alpha x} dx = \\ &= V^N \left(\prod_{i=1}^N m_i \right)^{3/2} (2\pi)^{3N/2} \alpha^{-3N/2}. \end{aligned}$$

Se E è l'energia totale del sistema

$$-\frac{d}{d\alpha} \log \Phi(\alpha) \Big|_{\alpha=\theta} = E$$

da cui

$$\frac{3N}{2\theta} = E \implies \theta = \frac{3N}{2E}.$$

Conseguentemente

$$\Phi(\theta) = V^N \left(\prod_{i=1}^N m_i \right)^{3/2} (2\pi)^{3N/2} \frac{(2E)^{3N/2}}{(3N)^{3N/2}}$$

Per B troviamo invece

$$B = \frac{d^2}{d\alpha^2} \log \Phi(\alpha) \Big|_{\alpha=\theta} = \frac{3N}{2\theta^2} = \frac{3N}{2} \frac{4E^2}{9N^2} = \frac{2E^2}{3N}$$

e così abbiamo l'espressione asintotica per $\Omega(E)$,

$$\begin{aligned} \Omega_{\text{asym}}(E) &= \frac{\Phi(\theta) e^{\theta E}}{\sqrt{2\pi B}} = V^N \left(\prod_{i=1}^N m_i \right)^{3/2} (2\pi)^{3N/2} e^{3N/2} \frac{(2E)^{3N/2}}{(3N)^{3N/2}} \frac{1}{(2\pi)^{1/2}} \frac{(3N)^{1/2}}{(2)^{1/2} E} = \\ &= V^N \left(\prod_{i=1}^N m_i \right)^{3/2} \frac{(2\pi e)^{3N/2}}{(3N/2)^{3N/2} (2\pi(3N)/2)^{1/2}} \frac{3N}{2} E^{(3N/2)-1} \end{aligned}$$

che va confrontata con la formula esatta

$$\Omega(E) = V^N \left(\prod_{i=1}^N m_i \right)^{3/2} \frac{(2\pi)^{3N/2}}{\Gamma((3N/2) + 1)} \frac{3N}{2} E^{(3N/2)-1}$$

Come si vede l'unica variazione sta nella sostituzione della funzione Γ di Eulero con la sua formula asintotica di Sterling.

Nel prossimo capitolo torneremo ad occuparci dei gas ideali.

II.2.6 Il teorema dell'equipartizione dell'energia

In questa sottosezione ricaveremo un teorema molto importante nella fisica statistica e nella termodinamica: il teorema di **equipartizione dell'energia**.

Un caso particolare del teorema di equipartizione si ottiene, come noto, considerando il gas ideale monoatomico (descritto nella sottosezione precedente). La distribuzione dell'energia di

una singola molecola vale approssimativamente $u_i^{(\theta)}$, perciò, nel caso del gas ideale, essendo

$$\begin{aligned}\phi_i(\theta) &= V(2\pi m_i)^{3/2} \frac{1}{\theta^{3/2}} \\ \omega_i(x) &= \frac{V(2\pi m_i)^{3/2}}{\Gamma(3/2)} x^{1/2} = 2\pi V(2m_i)^{3/2} x^{1/2}\end{aligned}$$

di modo che

$$u_i^{(\theta)}(x) = \frac{2}{\pi^{1/2}} \theta^{3/2} x^{1/2} e^{-\theta x}$$

da cui

$$\langle H_i \rangle \simeq \int x u_i^{(\theta)}(x) dx = -\frac{d}{d\theta} \log \phi_i = \frac{3}{2\theta}$$

Come si vede a prescindere dalle masse che abbiamo ammesso poter essere diverse, tutte le molecole hanno la medesima energia media identificata in modo univoco dal parametro θ . Questo risultato, che associa la medesima energia media a tutte le particelle anche diverse ma con lo stesso numero di gradi di libertà, discende, appunto, da un fatto più generale, il teorema di equipartizione, che andiamo adesso ad enunciare e dimostrare.

Consideriamo un sistema G le cui componenti G_i posseggano ℓ gradi di libertà. Assumiamo che la hamiltoniana del sottosistema G_i sia data solo dalla energia cinetica e perciò sia una forma quadratica nelle variabili di impulso (forma a coefficienti dipendenti eventualmente dalle \mathbf{q}_i). Inoltre, ogni componente ammetta la stessa hamiltoniana. Allora

$$V_i(x) = \int_{H(\mathbf{x}_i) < x} d^{2\ell} \mathbf{x}_i = \int d^\ell \mathbf{q}_i \int_{H(\mathbf{x}_i) < x} d^\ell \mathbf{p}_i$$

dove il secondo dominio di integrazione è la parte interna di un'ellissoide nello spazio degli impulsi (essendo le \mathbf{q}_i fissate). Ne viene che il secondo integrale sarà proporzionale a $x^{\ell/2}$ e con ciò

$$V_i(x) = x^{\ell/2} \int d^\ell \mathbf{q}_i \psi(\mathbf{q}_i) = c_1 x^{\ell/2}.$$

Ora,

$$\begin{aligned}\Omega_i(x) &= c_2 x^{(\ell/2)-1} \\ \Phi_i(\theta) &= c_2 \int_0^{+\infty} x^{(\ell/2)-1} e^{-\theta x} dx = c_2 \theta^{-\ell/2} \int_0^{+\infty} y^{(\ell/2)-1} e^{-y} dy = c_3 \theta^{-\ell/2}.\end{aligned}$$

conseguentemente

$$\langle H_i \rangle \simeq -\frac{d}{d\theta} \log \Phi_i = \frac{\ell}{2\theta}$$

Ogni grado di libertà (in cui la hamiltoniana risulti quadratica) porta un contributo $1/2\theta$ all'energia media.

Il risultato che abbiamo ottenuto è valido a parte la solita correzione $o(1/n)$, tuttavia è possibile vedere la sua versione esatta. Nel caso in cui H è quadratica negli impulsi, dal teorema di Eulero, abbiamo

$$H = \frac{1}{2} \sum_{k=1}^{\ell} p_k \frac{\partial H}{\partial p_k}$$

e perciò

$$\begin{aligned}\langle H_i \rangle &= \frac{1}{\Omega(E_{\text{tot}})} \int_{\Sigma_{E_{\text{tot}}}} \frac{H d\Sigma}{|\nabla H_{\text{tot}}|} = \frac{1}{2\Omega(E_{\text{tot}})} \sum_{k=1}^{\ell} \int_{\Sigma_{E_{\text{tot}}}} p_k \frac{\partial H}{\partial p_k} \frac{d\Sigma}{|\nabla H_{\text{tot}}|} = \\ &= \frac{1}{2\Omega(E_{\text{tot}})} \sum_{k=1}^{\ell} \int_{\Sigma_{E_{\text{tot}}}} p_k \frac{\partial H_{\text{tot}}}{\partial p_k} \frac{d\Sigma}{|\nabla H_{\text{tot}}|}\end{aligned}$$

Poiché

$$\frac{\partial H_{\text{tot}}}{\partial p_k} \frac{1}{|\nabla H_{\text{tot}}|} = (\hat{\mathbf{n}})_k$$

dove $\hat{\mathbf{n}}$ è il versore normale alla superficie Σ_E , abbiamo

$$\sum_{k=1}^{\ell} \int_{\Sigma_{E_{\text{tot}}}} p_k \frac{\partial H_{\text{tot}}}{\partial p_k} \frac{d\Sigma}{|\nabla H_{\text{tot}}|} = \int_{\Sigma_{E_{\text{tot}}}} \mathbf{p}_i \cdot \hat{\mathbf{n}} d\Sigma$$

dove \mathbf{p}_i è il vettore impulso di G_i allungato in \mathbb{R}^{2s} . Dal teorema di Gauß,

$$\int_{V(E_{\text{tot}})} \text{div } \mathbf{p}_i d^{2s} \mathbf{X} = \sum_{k=1}^{\ell} \int_{V(E_{\text{tot}})} \frac{\partial p_k}{\partial p_k} d^{2s} \mathbf{X} = \ell V(E_{\text{tot}})$$

sicch 

$$\langle H_i \rangle = \frac{\ell V(E_{\text{tot}})}{2\Omega(E_{\text{tot}})}$$

Dal confronto con la formula approssimata troviamo

$$\theta = \frac{\Omega(E)}{V(E)} + o\left(\frac{1}{n}\right) = \frac{d}{dE} \log V(E) + o\left(\frac{1}{n}\right)$$

II.3 L'insieme canonico di Gibbs

All'ultimo argomento di questo capitolo abbiamo voluto dare la dignit  di nuova sezione per l'importanza del suo contenuto. In realt , si tratta ancora di una applicazione dei metodi statistici introdotti, ma ci consente di introdurre un concetto fisico fondamentale: l'**equilibrio termico**. In questo ambito ricaveremo la ben nota legge di distribuzione canonica di Gibbs. Ancora, come per la legge di Boltzmann, si faccia attenzione alla grande mole di lavoro che abbiamo dovuto compiere per arrivare a una giustificazione su basi rigorose e critiche dell'insieme statistico di Gibbs.

II.3.1 Sistemi isolati e sistemi in equilibrio termico

Sistema isolato

Fino ad ora abbiamo considerato sistemi fisici G isolati, per i quali, cio , l'energia totale rappresentasse un integrale primo del moto. Evidentemente, si tratta di una idealizzazione, visto che qualsiasi sistema si utilizzi negli esperimenti di laboratorio, per quanto ben coibentato, interagisce con l'ambiente circostante. Tuttavia, quella di sistema isolato   una idealizzazione utile come approssimazione dei sistemi reali coibentati.

Sistema in equilibrio termico

Una idealizzazione di segno opposto, eppure altrettanto irrealizzabile in maniera rigorosa in laboratorio,   quella di bagno termico. In termini termodinamici, un bagno termico   una sorgente infinita di calore: per quanto calore possa cedere esso mantiene costante la sua temperatura, perci  presenta capacit  termica infinita. Un sistema G in contatto con un bagno termico, raggiunge necessariamente all'equilibrio termico la temperatura del bagno termico che, invece, pu  essere ritenuto come un sistema isolato. In pratica un sistema G in contatto con un bagno termico G^* pu  essere riguardato come una componente molto piccola di G^* .

Sistema microcanonico e sistema canonico

Perci , se nel caso isolato, la legge di distribuzione   data da

$$\frac{1}{\Omega(E) |\nabla H|}$$

ed   definita sulla superficie $H = E$, nel caso di equilibrio termico con una sorgente infinita, la legge di distribuzione si estende su tutto lo spazio Γ (e non su una superficie Σ_E essendo l'energia variabile) e vale, come ricavato nella sezione precedente,

$$\rho(\mathbf{x}) \equiv \frac{e^{-\theta^* H(\mathbf{x})}}{\Phi(\theta^*)} \quad (\text{II.5})$$

dove θ^*   l'unica radice dell'equazione

$$-\frac{d}{d\alpha} \log \Phi^*(\alpha) = E^*$$

(le quantit  con uno star all'indice si riferiscono al sistema infinito G^*).

L'idealizzazione di sistema isolato reca alla cosiddetta distribuzione **microcanonica** su Σ_E , mentre l'idealizzazione di sistema all'equilibrio con un bagno infinito, porta alla distribuzione (II.5) su Γ che si dice **canonica** o di Gibbs.

II.3.2 Distribuzione canonica

Soffermiamoci ancora un attimo sulla distribuzione canonica, $\mathbf{x} \in \Gamma \subset \mathbb{R}^{2s}$,

$$\frac{e^{-\theta^* H(\mathbf{x})}}{\Phi(\theta^*)}$$

Questa porta alla distribuzione in energia data da

$$\frac{\Omega(H) e^{-\theta^* H}}{\Phi(\theta^*)}$$

di modo che

$$\langle H \rangle = \frac{1}{\Phi(\theta^*)} \int H \Omega(H) e^{-\theta^* H} = - \left. \frac{d}{d\alpha} \log \Phi(\alpha) \right|_{\alpha=\theta^*}$$

Questo mostra che il ruolo di $\langle H \rangle$ è analogo quello di E nel sistema microcanonico, così come il ruolo di θ^* corrisponde a quello di θ . In particolare θ^* , essendo l'unico parametro veramente dipendente dal sistema termalizzante, deve essere legato alla temperatura.

Consideriamo ora due componenti entro il sistema in equilibrio termico G , G_1 e G_2 . Vogliamo ricavare la legge di distribuzione entro lo spazio delle fasi Γ_1 del sistema G_1 . Se $M \subset \Gamma_1$, abbiamo

$$\int_M \rho_{G_1}(\mathbf{x}_1) d^{2s_1} \mathbf{x}_1 \equiv \int_{M \times \Gamma_2} \frac{e^{-\theta^* (H_1(\mathbf{x}_1) + H_2(\mathbf{x}_2))}}{\Phi(\theta^*)} d^{2s} \mathbf{x} = \int_M d^{2s_1} \mathbf{x}_1 e^{-\theta^* H_1(\mathbf{x}_1)} \int_{\Gamma_2} d^{2s_2} \mathbf{x}_2 \frac{e^{-\theta^* H_2(\mathbf{x}_2)}}{\Phi(\theta^*)}$$

Perciò la distribuzione nello spazio delle fasi Γ_1 vale

$$\rho_{G_1}(\mathbf{x}_1) = e^{-\theta^* H_1(\mathbf{x}_1)} \int_{\Gamma_2} d^{2s_2} \mathbf{x}_2 \frac{e^{-\theta^* H_2(\mathbf{x}_2)}}{\Phi(\theta^*)}$$

Siccome, per quanto ottenuto in precedenza,

$$\int_{\Gamma_2} e^{-\theta^* H_2} d^{2s_2} \mathbf{x}_2 = \Phi_2(\theta^*)$$

quindi,

$$\rho_{G_1}(\mathbf{x}_1) = \frac{e^{-\theta^* H_1(\mathbf{x}_1)}}{\Phi_1(\theta^*)} \tag{II.6}$$

Analogamente avremo per ρ_{G_2} , sicché troviamo la notevole legge di composizione

$$\rho_G = \rho_{G_1} \rho_{G_2}.$$

Si ottiene così che, a differenza che nell'insieme microcanonico, *le componenti di un sistema canonico sono mutuamente indipendenti*. Evidentemente, nell'insieme microcanonico così non poteva essere a causa del vincolo di conservazione dell'energia.

**Relazione
tra insieme
microcanonico e
insieme canonico**

Un fatto importante è che la formula (II.6) per la distribuzione di una componente (non importa quanto grande) di un sistema canonico coincide con la legge di Boltzmann per la distribuzione di una componente **piccola** di un insieme microcanonico.

Ora, chiaramente, i calcoli eseguiti con la distribuzione canonica sono enormemente più semplici di quelli per la microcanonica, visto che è più facile lavorare con quantità aleatorie indipendenti. La domanda che si pone è quanto sia buona l'approssimazione consistente nell'usare la distribuzione canonica per calcoli concernenti sistemi microcanonici.

Abbiamo già visto che le funzioni di fase che ci interessano sono quelle additive. Le loro medie si ottengono come somme delle medie sulle singole molecole: queste possono essere calcolate usando la distribuzione canonica, perciò, l'approssimazione, di calcolare i valori medi per le funzioni di fase microcanoniche utilizzando la distribuzione canonica è buona. Tuttavia, le dispersioni non possono essere calcolate tramite la canonica, altrimenti riterremo i vari addendi scorrelati cosa che nell'insieme microcanonico è falsa.

Come esempio degli errori che si possono fare usando la distribuzione di Gibbs al posto della microcanonica, si pensi all'**energia**. Essa ha, rispetto alla ρ di Gibbs dispersione finita, ma un insieme microcanonico è zero!!

Infine, si noti l'economia concettuale che si è seguita per giungere agli insiemi microcanonico e canonico: entrambi discendono dal teorema di Liouville e da una discussione di carattere ergodico, inevitabilmente. È sbagliato assumere la distribuzione di Gibbs in modo assiomatico,

indipendentemente dall'insieme microcanonico.

I metodi della meccanica statistica

La meccanica statistica si basa sul metodo del calcolo delle medie di fase per calcolare i risultati delle misure sperimentali delle osservabili. In questa sede, ci si chiede per quale motivo questa strategia è così potente dal punto di vista teorico. In altri termini si vuole una giustificazione su basi meccaniche della validità della meccanica statistica.

Il problema è, oggi, in alcune sue parti ancora aperto, perciò questo rapido resoconto non può essere più di tanto approfondito. Lo scopo è quello di rendere plausibili gli argomenti della meccanica gibbsiana attraverso una rassegna delle sue giustificazioni da Birkhoff fino a Malament, Zabell e Vranas. Quanto scritto è frutto dello studio degli articoli sui fondamenti della meccanica statistica di Janneke van Lith.

III.1 Impostazione del problema

La meccanica statistica di Gibbs e la sua giustificazione

La questione che vogliamo indagare è per quale motivo la meccanica statistica di Gibbs è corretta da un punto di vista teorico, cioè in quale senso la meccanica microscopica dettata dalle leggi di Newton rende conto degli aspetti termodinamici della materia. Lo scopo che ci poniamo non è, in ogni caso, quello di dare una riduzione della termodinamica alla meccanica statistica, ma spiegare le origini microscopiche dei metodi della meccanica statistica. Il fatto che la meccanica statistica sia in grado di spiegare i fenomeni osservati sperimentalmente e, con essi, i principi della termodinamica, una volta che si sia operato un ponte tra le grandezze macroscopiche e le osservabili della meccanica statistica stessa, non verrà qui indagato.

Quello che vogliamo spiegare, o quanto meno rendere plausibile, è il fatto che per certi sistemi fisici isolati e all'equilibrio, il valore di una osservabile termodinamica può essere dedotto con successo calcolando la media microcanonica della funzione di fase corrispondente.

I metodi della meccanica gibbsiana

Il metodo del calcolo della media microcanonica di fase è centrale nella meccanica statistica. L'insieme canonico e quello grancanonico si ottengono infatti a partire da quello microcanonico, perciò una volta spiegato il motivo del successo della strategia di media microcanonica, avremo una base fondante per l'intera meccanica statistica.

La misura microcanonica

Ricordiamo, prima di passare in rassegna i vari tentativi di spiegazione della strategia di Gibbs, che la misura microcanonica è la seguente

$$d\mu(\mathbf{x}) \equiv \delta(H(\mathbf{x}) - E) d^{2\ell}\mathbf{x}$$

per un sistema hamiltoniano isolato H , avente energia E e possedente ℓ gradi di libertà. Per comodità abbiamo inteso $\mathbf{x} = (\mathbf{p}, \mathbf{q})$.

Come si vede la misura microcanonica μ è una misura sulla varietà $H(\mathbf{x}) = E$ che, per il teorema di Liouville e per la conservazione dell'energia, è invariante sotto il flusso canonico T_t generato dalle equazioni di Hamilton. In generale, se Σ_E indica la varietà ad energia E , abbiamo che $(\Sigma_E, \mathcal{B}, \mu, T_t)$ è un sistema dinamico continuo dotato di misura invariante.

Se con f_{mis} indichiamo il valore sperimentale di una osservabile termodinamica alla quale sia associata una funzione f sullo spazio delle fasi di cui sopra, si tratta di spiegare per quale motivo

$$f_{\text{mis}} = \frac{1}{\Omega(E)} \int_{\Sigma_E} f(\mathbf{x}) d^{2\ell}\mu(\mathbf{x})$$

dove

$$\Omega(E) = \int_{\Sigma_E} d^{2\ell} \mu(\mathbf{x})$$

è la funzione di struttura del sistema.

**Ingredienti per
la spiegazione
dei metodi
di Gibbs**

Una spiegazione ragionevole del metodo di Gibbs, dovrebbe tener conto dei seguenti aspetti che caratterizzano i sistemi statistici

- (i) equilibrio termodinamico;
- (ii) alto numero di particelle costituenti il sistema;
- (iii) proprietà dinamiche;

III.2 L'approccio ergodico

Medie nel tempo

Per risolvere il problema si tratta di dare una formalizzazione matematica alla funzione f_{mis} . Siccome la misura di f avviene in un certo intervallo di tempo T , si può ritenere plausibile che f_{mis} sia la media temporale della f sulla traiettoria, cioè

$$f_{\text{mis}} = \frac{1}{T} \int_0^T f(T_t \mathbf{x}) dt,$$

dove \mathbf{x} è il punto dello spazio delle fasi (microstato) occupato dal sistema all'inizio della misurazione. La definizione data di f_{mis} appare molto problematica per varie ragioni. In primo luogo, è plausibile ma alquanto difficile da giustificare che f_{mis} sia la media temporale della f , tutto ciò che si può dire con tranquillità al riguardo è semmai che

$$\inf_{t \leq T} f(T_t \mathbf{x}) \leq f_{\text{mis}} \leq \sup_{t \leq T} f(T_t \mathbf{x}).$$

In ogni caso, la f_{mis} dipende, a priori, in modo forte dal dato iniziale \mathbf{x} del quale, sperimentalmente non sappiamo nulla e che, anzi, possiamo aspettarci vari in modo molto forte al variare della scelta dell'istante iniziale. Sperimentalmente, però, il sistema è omogeneo nel tempo, cioè la misura non dipende in nessun modo (se il sistema è ben isolato) dall'istante iniziale. Un modo per alleviare i guai introdotti dalla presenza del termine T che non è noto e che dipende dall'interazione dello strumento di misura con il sistema, è quello di prendere il limite per $T \rightarrow \infty$ nella prima definizione di f_{mis} . Una giustificazione di questa procedura di limite, può risiedere nel fatto sperimentale (si pensi al moto browniano) che i tempi caratteristici per l'evoluzione del sistema sono molto più brevi del tempo T che occorre allo strumento per interrogare il sistema stesso. Dal momento che perfino la prima assunzione (quella di media nel tempo) era molto vaga, l'idea che possa effettivamente essere

$$f_{\text{mis}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(T_t \mathbf{x}) dt \equiv \bar{f}_{\mathbf{x}}$$

risulta assai poco chiara.

**Teorema di
Birkhoff e
suoi corollari**

Tuttavia, l'approccio ergodico al problema della giustificazione del metodo di Gibbs si fonda su questa assunzione e perciò non la discuteremo oltre. Il fatto importante è che $\bar{f}_{\mathbf{x}}$ esiste quasi per ogni dato iniziale (rispetto alla misura microcanonica) grazie al teorema di Birkhoff, che si basa sull'invarianza della misura microcanonica garantita dal teorema di Liouville. L'approccio ergodico standard alla nostra questione, ci obbliga adesso a dimostrare che

$$\bar{f}_{\mathbf{x}} = \langle f \rangle \equiv \frac{1}{\Omega(E)} \int_{\Sigma_E} f(\mathbf{x}) d^{2\ell} \mu(\mathbf{x})$$

visto che il primo membro eguaglia f_{mis} .

Il primo corollario al teorema di Birkhoff asserisce che l'eguaglianza ipotizzata è effettivamente valida, per quasi ogni \mathbf{x} , se il sistema è **metricamente indecomponibile**. Quest'ultima ipotesi si verifica se il flusso T_t è tale che non esiste una partizione di Σ_E in due sottoinsiemi misurabili, invarianti, di misura positiva. Correntemente, l'indecomponibilità metrica viene chiamata anche **transitività metrica** o, più semplicemente, **ergodicità**.

**I teoremi
ergodici e il
teorema KAM**

Apparentemente, il nostro problema sarebbe risolto se ci restringessimo a sistemi ergodici. Il

modello di Maxwell e Boltzmann di gas ideale è uno dei pochi esempi noti di sistemi ergodici (solo la dimostrazione – dovuta a Sinai – di questo fatto richiederebbe qualche centinaio di pagine di matematica raffinata). Per di più il teorema KAM mostra un modo canonico per produrre controesempi di sistemi meccanici non ergodici, la qualcosa appare come una pesante limitazione dell'approccio ergodico standard.

Altri problemi
dell'approccio
ergodico

Tuttavia, quandanche ci si restringa a considerare sistemi metricamente indecomponibili, restano due gravi problemi a screditare la bontà della spiegazione in termini ergodici: per primo l'identificazione del valore sperimentale con la media su tempi infiniti (praticamente priva di senso, finché non si fornisce la velocità con cui la media raggiunge il suo limite), per secondo il problema di come trattare gli insiemi di misura nulla di punti per cui neppure il teorema di Birkhoff vale. Infatti, il comportamento di un insieme di misura microcanonica nulla può essere molto importante dal punto di vista dinamico. Consideriamo per esempio una famiglia di traiettorie per cui la media temporale non esista o sia molto diversa dalla media di fase: ebbene essa rappresenta un insieme di misura nulla! Appare così molto importante poter controllare il comportamento dei punti appartenenti agli insiemi esclusi dal teorema di Birkhoff e questo è del tutto impossibile coi metodi ergodici descritti.

Contraddizioni
interne al
metodo ergodico

C'è di più, il primo corollario al teorema di Birkhoff sostiene che se il sistema ammette medie di fase eguali alle medie temporali (quasi per ogni traiettoria), allora è metricamente indecomponibile. Ne viene che la strategia ergodica reca al risultato che solo per i sistemi ergodici il metodo di Gibbs è valido. La cosa è falsa, visto che sono noti esempi di sistemi altamente non ergodici (la catena di oscillatori) per i quali la meccanica statistica riproduce i fatti sperimentali (la catena di oscillatori assume comportamento gibbsiano quando il numero degli oscillatori diviene considerevole).

La chiave di questo paradosso, sta probabilmente nell'identificazione di f_{mis} con $\bar{f}_{\mathbf{x}}$ che è sbagliata, ma allora sorge il problema di fornire una nuova interpretazione per f_{mis} .

Mancanza del
riferimento
al numero di
gradi di libertà

Un ultimo punto critico, per quanto riguarda la teoria ergodica dei sistemi statistici, risiede nel fatto che in nessun modo si fa riferimento all'alto numero di gradi di libertà del sistema che è, invece, una proprietà essenziale (si pensi proprio alla catena di oscillatori di cui sopra).

III.3 L'approccio di Khinchin

Originalità
del lavoro
di Khinchin

Il metodo usato da I. A. Khinchin per attaccare il problema della fondazione della meccanica statistica, sta in mezzo tra l'approccio ergodico e la teoria del limite termodinamico (della quale qui non ci occuperemo). Nel suo *Mathematical Foundations of Statistical Mechanics* (1949, Dover, new York), Khinchin delinea prima l'approccio ergodico così come descritto nella sezione precedente e poi innova la teoria facendo uso del teorema del limite centrale. Sebbene molto elegante e compatta la teoria di Khinchin ha validità limitata e, soprattutto, è inficiata dalla presenza di tre punti oscuri uno dei quali difficilmente superabile (tanto che l'uso del teorema del limite centrale nasce e muore con questo testo). Nonostante questo, molte (se non tutte) le idee successive per il superamento delle difficoltà legate all'esplicazione del metodo di Gibbs, sono più o meno esplicitamente contenute nell'opera dell'autore russo.

Integrali primi
ed ergodicità

Il primo problema della trattazione è quello della identificazione del risultato sperimentale f_{mis} con la media temporale $\bar{f}_{\mathbf{x}}$. Il secondo problema (il più grave) è legato, sembrerà strano, a quella che è senza dubbio la più potente trovata di Khinchin: la riduzione alla considerazione delle funzioni additive. Vediamo questo punto in un qualche dettaglio. Il problema della ergodicità di un sistema si presenta quando, oltre all'energia si presentano altri integrali primi del moto. Supponiamo che L sia un integrale primo misurabile indipendente da H . Sulla varietà a energia costante, si potrà certo trovare un numero α tale che gli insiemi

$$L_1 \equiv \{\mathbf{x} \in \Sigma_E \mid L(\mathbf{x}) < \alpha\}$$

$$L_2 \equiv \{\mathbf{x} \in \Sigma_E \mid L(\mathbf{x}) \geq \alpha\}$$

abbiano misura positiva e partiscano Σ_E . Allora, la sola presenza di L compromette l'indecomponibilità metrica e con ciò l'approccio ergodico standard al problema.

Riduzione
dell'algebra
delle osservabili

Khinchin suggerisce come evitare questa semplice obiezione alla indecomponibilità. Egli, infatti, si limita a considerare un'algebra ristretta di osservabili, di modo da ridurre al minimo

i possibili integrali primi indipendenti da H . Ridurre l'algebra delle osservabili, significa ridurre l'algebra dei misurabili in maniera coerente, e perciò rendere più difficile la decomposizione metrica.

Paradosso metodologico

Il problema è che Khinchin opera il "taglio" delle osservabili avendo in mente i gas ideali e perciò finendo con il limitare troppo la forza della sua teoria. La riduzione è, infatti, alle funzioni additive, cioè a quelle funzioni di fase che si esprimono come somma di funzioni a valori negli spazi delle fasi delle singole particelle. La cosa non sarebbe poi troppo grave se non si richiedesse che anche l'hamiltoniana dovesse rispondere a questo requisito. Rendere additiva l'hamiltoniana significa trascurare le interazioni tra le particelle, perciò la possibilità stessa di una termalizzazione. La termodinamica sembra impossibile da riprodurre. Khinchin è cosciente di questo guaio che si viene a creare nella sua speculazione e tenta di rimediare facendo appello al cosiddetto **paradosso metodologico**: i termini di interazione ci sono e garantiscono la termalizzazione e l'esistenza stessa di una termodinamica, ma nel calcolo delle medie sono piccoli e perciò trascurabili. Come si vede il ragionamento fila per i gas perfetti o i reali diluiti, ma cade per i liquidi e per le transizioni di fase: eppure la meccanica statistica funziona anche in quelle condizioni!!

Uso dell'additività della hamiltoniana

Il lettore attento si sarà chiesto perché richiedere che anche l'hamiltoniana sia additiva, visto che stiamo ragionando sulla varietà in cui essa è fissata e perciò non ce ne occorrono medie? Il fatto è che l'apparato del teorema del limite centrale che consente a Khinchin di ottenere la distribuzione di Gibbs, funziona solo se l'hamiltoniana è separata in somma di hamiltoniane di singole particelle. Ma il teorema del limite centrale non è importante solo perché consente di ritrovare la distribuzione canonica e quella di Boltzmann su basi matematicamente fondate, esso permette anche di riformulare il teorema ergodico in modo da non essere più costretti a contare sull'ipotesi di transitività metrica.

Medie delle funzioni additive

Infatti, per le funzioni additive, Khinchin mostra che le regioni di Σ_E in cui le osservabili differiscono poco dalla loro media di fase, hanno misura microcanonica molto grande, per numero di particelle N molto grande. In altre parole, con grande probabilità microcanonica, le osservabili hanno valori vicini a quelli delle loro medie. Khinchin riesce pure a mostrare che questo implica che le medie temporali sono vicine a quelle di fase con grande probabilità microcanonica quando N è grande.

Abbandono dell'ipotesi ergodica

Ora, il successo più notevole di Khinchin sta nel render conto del metodo gibbsiano senza usare l'ergodicità, ma riferendosi al fatto che i sistemi termodinamici sono costituiti da un numero impressionante di sottosistemi. Tuttavia, e qui compare il terzo problema, restano gli insiemi di piccola misura microcanonica di cui non si riesce a dire nulla. Se, nell'approccio ergodico, era importante ma impossibile controllare il comportamento per gli insiemi di misura nulla, nell'approccio di Khinchin il problema diviene addirittura più drammatico.

Pregi e difetti della teoria di Khinchin

Riassumendo, il metodo di Khinchin è interessante perché più aderente alla realtà fisica della termodinamica, importanza del fatto che N è grande, interesse ai sistemi non ergodici, restringimento della classe delle osservabili, ma risulta indebolito dal paradosso metodologico e dal problema degli insiemi di misura piccola di punti per i quali le osservabili sono a media spaziale lontana. Inoltre, la teoria è valida solo per i gas debolmente interagenti e non spiega, per esempio, le transizioni di fase.

Teorema di dispersione di Khinchin-Lanford

In ogni caso, il teorema di dispersione di Khinchin circa il comportamento delle osservabili rispetto alla media è stato migliorato nell'ambito della teoria del limite termodinamico (Lanford, 1973) in modo da estenderlo a funzioni non additive (ma soggette ad ipotesi più ragionevoli). Questo fatto ci sarà utile tra poco.

Integrali primi additivi (Landau)

Prima di concludere, un cenno sulla possibilità di ridurre la teoria di Khinchin a quella ergodica standard per quanto riguarda l'identificazione di medie spaziali e temporali. Se nel considerare i sottosistemi, possiamo giungere alle singole particelle, siccome esse sono regolate da una hamiltoniana di particella libera, ciascuna ammette come integrali del moto energia, momento angolare e quantità di moto (visto che i gradi di libertà sono sei, ogni altro integrale dipende da questi). Ne viene che per l'intero sistema gli integrali primi **additivi** e **indipendenti** sono al più energia totale, momento angolare totale e impulso. Se il gas

è confinato in una cella, l'impulso perde il suo carattere di integrale. Se la cella non è sferica o cilindrica (ma a quel punto basterà ragionare sull'intersezione della varietà e a energia costante con quella a momento costante), resta come unico integrale additivo l'energia. Considerando solo osservabili additive, troviamo l'indecomponibilità metrica in senso esteso del nostro sistema.

Anche in questo modo, non si risolve il problema di misura zero.

III.4 La strategia di Malament, Zabell e Vranas

Il tentativo meglio riuscito nella spiegazione del metodo gibssiano si deve a David Malament e Sandy Zabell (1980) e a Peter Vranas (1998).

Cambiamento
del punto
d'avvio
della teoria

Il punto di partenza sta nell'introdurre una misura di probabilità *reale* che a differenza della misura microcanonica non sia un semplice artificio matematico. Il fatto è che negli approcci precedenti la μ nasceva dalla teoria ergodica e, in ultima analisi, dall'assunzione che $f_{\text{mis}} = \bar{f}_{\mathbf{x}}$, mentre Malament e Zabell scelgono premesse diverse di modo che l'interpretazione di f_{mis} nasca autonomamente e non sia dettata fin dall'inizio.

Per cominciare ci concentriamo sui contenuti dell'articolo di Malament e Zabell, solo dopo vedremo le correzioni apportate da Vranas. Considerando il complesso si ha che a fare con una teoria piuttosto solida, alla quale le cui ipotesi richiedono pochi chiarimenti e possono essere assunte senza troppa fatica come postulati, senza dover forzare fatti fisici evidenti, come avveniva quando si doveva supporre l'indecomponibilità metrica.

I due ingredienti
della teoria
di Malament
e Zabell

In estrema sintesi la teoria di Malament e Zabell si basa sul teorema di dispersione di Khinchin e Lanford di cui nella sezione precedente e sul fatto che nello spazio delle fasi esiste una effettiva distribuzione di probabilità $P_{\textcircled{a}}$ (il pedice @ sta per *actual*) che coincide con μ una volta che si accettino determinate ipotesi su $P_{\textcircled{a}}$ stessa.

Esistenza
"fisica" della
distribuzione $P_{\textcircled{a}}$

Può apparire sconvolgente l'introduzione di una distribuzione di probabilità non derivante da un artificio matematico, ma assunta avere un significato fisico. Il problema del significato dell'esistenza fisica di $P_{\textcircled{a}}$ è troppo complicato da affrontare nella sua interezza perché richiede la conoscenza delle varie interpretazioni della teoria delle probabilità di Kolmogorov (interpretazione di von Mises, di Popper, etc.) Qui possiamo limitarci a rendere plausibile l'esistenza di $P_{\textcircled{a}}$. La cosa che dà fastidio nel dover introdurre una $P_{\textcircled{a}}$ è che si sta trattando un problema deterministico, per cui non sembra ci sia alcuna ragione per dover parlare di probabilità. Le cose, però, non stanno in questi termini. Le leggi di Newton sono inapplicabili nel dominio della meccanica statistica: esse ci dicono qualcosa solo se sappiamo specificare il dato iniziale, ma questo è sperimentalmente impossibile, visto che dovremmo conoscere a un istante fissato una $2N$ -upla di numeri, dove N è dell'ordine di 10^{24} . Ne viene che, anche a causa della mera imprecisione delle misure (che si avrebbe pure per $N = 1$) il dato iniziale ci è in parte sconosciuto e tutta l'informazione sperimentale che abbiamo si formalizza in una distribuzione di probabilità. Stiamo così giustificando l'esistenza di $P_{\textcircled{a}}$ usando il punto di vista intersoggettivista che è, a mio avviso, il più indicato a descrivere una situazione sperimentale. Per concludere vogliamo abbattere ogni ulteriore scetticismo riferendoci al lancio di una moneta. Nessuno dubita che si tratta di un fenomeno deterministico, tuttavia, a causa della nostra ignoranza circa il dato iniziale e circa la dinamica dell'atterraggio, siamo indotti in modo del tutto naturale ad associare probabilità $1/2$ all'evento "testa" come all'evento "croce". In meccanica statistica, come in generale in ogni processo in cui le variabili in gioco sono troppo più numerose delle quantità che possiamo controllare, si ha bisogno di introdurre una distribuzione di probabilità. A posteriori, solo introducendo la $P_{\textcircled{a}}$ siamo in grado di usare le altrimenti inutili equazioni di Newton.

Le ipotesi di
Malament e
Zabell su $P_{\textcircled{a}}$

Ebbene, Malament e Zabell ritengono di poter richiedere che $P_{\textcircled{a}}$ sia stazionaria e assolutamente continua rispetto alla misura microcanonica. Ora, se il sistema è metricamente indecomponibile, il secondo corollario al teorema ergodico, mostra che, sotto le ipotesi dette, $P_{\textcircled{a}} = \mu$. Usando il teorema di dispersione abbiamo che le funzioni di fase stanno sempre vicine alle medie spaziali con grande probabilità microcanonica. L'insieme su cui questo non avviene ha misura microcanonica piccola, e perciò probabilità effettiva $P_{\textcircled{a}}$ piccola. Questo risolve

il problema della misura piccola o nulla, visto che adesso parliamo di misura *reale*. Adesso importa poco definire f_{mis} come media, possiamo anzi assumere un punto di vista istantaneo nell'interpretazione di f_{mis} .

Uso dell'ergodicità

Quali sono i punti deboli di Malament e Zabell? In primo luogo, l'ipotesi di ergodicità che ci costringe a tornare ai problemi di cui nell'approccio standard; in secondo luogo l'ipotesi di assoluta continuità che, così come è posta sembra ingiustificabile.

Continuità per traslazione

Tuttavia, per quanto riguarda la seconda questione, Malament e Zabell dimostrano un teorema secondo cui l'assoluta continuità rispetto alla misura di Lebesgue (e, con un opportuna generalizzazione dovuta a Janneke van Lith, alla misura microcanonica su Σ_E) equivale alla continuità per traslazione. Quest'ultima proprietà si attribuisce alle misure per cui

$$\lim_{\mathbf{s} \rightarrow 0} \mu(d(A, \mathbf{s})) = \mu(A) \quad \forall A \in \mathcal{B}$$

dove

$$d(A, \mathbf{s}) = \{\mathbf{x} | \mathbf{x} - \mathbf{s} \in A\}.$$

In altri termini, una misura è continua per traslazione se dà probabilità vicine a insiemi traslati poco. Il fatto che $P_{\text{@}}$ goda della continuità per traslazione si può giustificare pensando alle imprecisioni che si hanno effettuando misure sul sistema: la probabilità di trovare un punto in A non può differire troppo dalla probabilità di trovarlo in un insieme vicino quale $d(A, \mathbf{s})$.

Il lavoro di P. B. M. Vranas: ε -ergodicità

Il lavoro di Vranas consente di superare il primo problema sorto nell'approccio di Malament e Zabell, cioè l'incomponibilità metrica. Anziché considerare sistemi ergodici, Vranas riesce a generalizzare i risultati dei predecessori a sistemi che siano ε -ergodici, cioè che ammettano un insieme invariante metricamente transitivo di misura $1 - \varepsilon$. A differenza dei sistemi strettamente transitivi, i sistemi ε -ergodici non prestano il fianco al teorema KAM se i tori invarianti hanno misura sufficientemente piccola. I vantaggi della formulazione ε -ergodica sono i seguenti

- (i) si allarga la classe di sistemi fisici a cui si può applicare la teoria;
- (ii) non si presta facilmente il fianco al teorema KAM;
- (iii) non esistono esempi di sistemi fisici per cui il metodo di Gibbs abbia successo e che siano chiaramente non ε -ergodici.

Come si vede, la strategia di Malament, Zabell e Vranas, risolve il problema della misura zero e delle ipotesi dinamiche troppo forti (ergodicità stretta), oltre a mettere fine alla questione sulla interpretazione di f_{mis} .

C'è da dire che molto ancora bisogna fare per rendere rigorosa (su basi meccaniche e di interpretazione probabilistica) l'affermazione secondo cui $P_{\text{@}}$ ha da essere continua per traslazione e per dimostrare la ε -ergodicità almeno per N grande dei sistemi di interesse fisico comune.

Riformulazione del concetto di equilibrio (J. van Lith)

Ma il vero valore dell'approccio descritto, sta nel fatto che, nella formulazione di Vranas e nella sua revisione di van Lith non è necessario richiedere la stretta stazionarietà di $P_{\text{@}}$. La caratterizzazione dell'equilibrio con misure strettamente stazionarie è, infatti, un paradosso del metodo gibbsiano: infatti, l'evoluzione hamiltoniana non consente di passare da distribuzioni dipendenti dal tempo a distribuzioni stazionarie, con ciò si vieterebbe il passaggio dal non equilibrio all'equilibrio che invece è fondamentale in termodinamica.

Questo significa che con Malament, Zabell e Vranas il programma della giustificazione del metodo delle medie microcanoniche è quasi concluso o che, almeno, esso ha raggiunto un grado di raffinatezza molto importante. Almeno per quanto mi riguarda, il livello di plausibilità delle tecniche di calcolo della meccanica statistica è adesso accettabile.