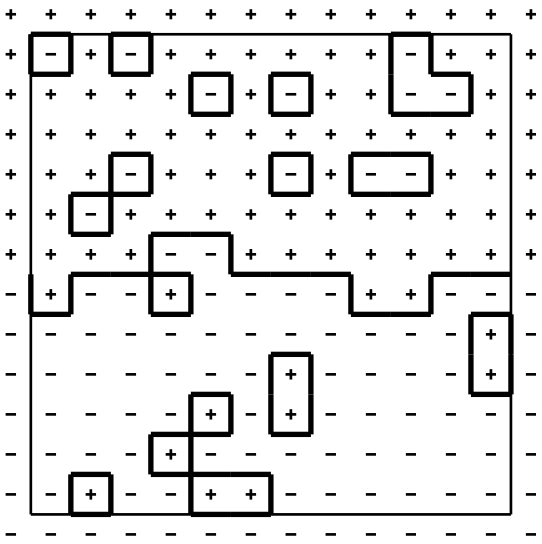
Giovanni Gallavotti

MECCANICA STATISTICA

Trattatello



Roma 1994

Trattatello di

Meccanica statistica

Giovanni Gallavotti
Dipartimento di Fisica
Università di Roma *La Sapienza*
00185 Roma

Roma 1994

A Daniela e Barbara

Introduzione

Raccolgo qui la collezione di manoscritti prestampa elaborati principalmente per la Enciclopedia Italiana delle Scienze Fisiche quale coordinatore della voce quadro Meccanica Statistica (con pochissime modifiche rispetto alle versioni stampate o in corso di stampa).*

Insieme forniscono un' analisi delle questioni fondamentali della Meccanica Statistica, soprattutto classica, elaborata nel corso di un ventennio durante il quale gli argomenti esposti sono stati oggetto di molti corsi e seminari in varie università. Alcuni manoscritti prestampa vennero elaborati per altri scopi (Dizionario del Novecento, o riviste scientifiche).

Non sempre la discussione è una pura riesposizione di idee ben stabilite; non è però il caso di fornire un elenco delle *novità*. Considererò un successo di questo lavoro se riceverò commenti (anche se sdegnati) sui punti che possono essere controversi: vorrà dire almeno che il lavoro avrà raggiunto l' obiettivo di sollevare una discussione sui fondamenti della Meccanica Statistica. Considero infatti di grande importanza una tale discussione e il parteciparvi sarà per me molto interessante.

Spero che questa collezione possa anche essere utile agli studenti, portando alla loro attenzione problemi ai quali nei corsi universitari vengono dedicati solo cenni, per *necessità di concretezza* (ossia perchè "queste cose non servono a nulla") o più semplicemente per mancanza di tempo. Non intendo però spingerli ad occuparsi di questioni di fondamenti. Io credo, anzi, che i più giovani *non* debbano occuparsi di questioni di fondamenti: a questi ci si può dedicare dopo che si è ottenuta una visione prospettica che solo la ricerca attiva e avanzata può fornire (o almeno il tentativo sofferto di svolgerla, nel corso di lunghi anni). Tanto più che chi si accosta troppo presto a tali problemi inevitabilmente cede alla tentazione di dedicare l' attenzione a questioni astratte, che non mi paiono mai rilevanti. In generale, infatti, non apprezzo studi sui fondamenti che non abbiano il carattere di concretezza che spero questi miei scritti abbiano, nonostante tutto.

Ma non voglio dire che gli studenti non debbano formarsi una visione "filosofica" dei problemi dell' area nella quale intendono lavorare. Perchè è

* in alcuni punti appaiono riferimenti ad altre voci dell' Enciclopedia, dovute spesso ad altri autori, o talvolta solo previste nel programma e non scritte. I riferimenti sono indicati con (v.), ovvero in forma più esplicita.

necessario, per produrre idee originali, o anche solo lavoro di buona qualità scientifica, che ognuno si formi convinzioni filosofiche precise sulla natura delle cose, pur nella consapevolezza che qualsiasi convinzione filosofica sui fondamenti (della Meccanica Statistica e della Fisica più in generale) per quanto chiara e irrefutabile possa apparire a chi la formula, anche se dopo lunghe e meditate vigilie, difficilmente potrà apparire meno che criticabile a qualunque altro interlocutore. E pur di essere, quindi, sempre disposti a rimettere tutto in discussione, ed evitare di credere di aver infine raggiunto la "verità", *una, immutabile e oggettiva* (nella quale è vano credere).

Desidero ringraziare l' Enciclopedia Italiana per aver stimolato la realizzazione di questo lavoro conferendomi il compito di coordinatore della sezione di Meccanica Statistica della Enciclopedia delle Scienze Fisiche. Ed è doveroso sottolineare che quest' opera è stata resa possibile dal supporto finanziario e soprattutto culturale dell' Istituto dell' Enciclopedia Italiana. Il clima creato nella redazione del Dizionario delle Scienze Fisiche, divenuto poi Enciclopedia, mi ha profondamente stimolato alla meditazione dei vari temi raccolti nel corso degli ultimi nove anni: ed esprimo qui la mia riconoscenza ai colleghi redattori. E' molto importante che in Italia ci siano ancora iniziative di questa natura: non immediatamente legate alla logica di profitto che muove un' editoria troppo preoccupata, al tempo stesso, di non correre rischi. Il contributo finanziario del CNR (Gruppo Nazionale di Fisica Matematica) ha poi reso possibile la stampa: ed esprimo al Consiglio Scientifico e ai professori C. Cercignani, C. Marchioro e V. Boffi la mia gratitudine per l' apprezzamento così mostrato per questo mio lavoro.

Ringrazio infine A. Alippi, G. Altarelli, P. Dominici e V. Cappelletti per essersi adoprati affinchè l' Enciclopedia Italiana potesse ancora una volta mostrare la sua straordinaria apertura culturale permettendo la riproduzione dei testi, dei quali detiene i diritti d' Autore, in questa raccolta (non destinata al commercio). Permettendone così una più facile circolazione negli ambienti scientifici italiani e stranieri che, con la loro influenza, hanno anche contribuito alla concezione e realizzazione. Questi testi sono riconoscibili dagli altri perchè recano nella prima pagina la dicitura "voce per l' Enciclopedia delle Scienze Fisiche" ovvero "per il Dizionario del Novecento".

La *Rivista del Nuovo Cimento* e il *Journal of Statistical Physics* hanno gentilmente concesso la riproduzione degli articoli ivi pubblicati.

Giovanni Gallavotti

Roma, dicembre 1994

Indice

Meccanica statistica classica 1

Insiemi statistici 35

Equipartizione e critica della meccanica statistica classica. 65

Modelli risolubili 87

Limite termodinamico 97

Moto browniano 109

Entropia e informazione 127

Serie del viriale 143

Transizioni di fase nel modello di Ising 149

Equazione di Boltzmann 191

Ergodicità e irreversibilità 205

Teoria ergodica 231

Caos 259

Voci lessicali 281

MECCANICA STATISTICA CLASSICA

per il "DIZIONARIO DELLE SCIENZE FISICHE"
edito dall' Istituto della Enciclopedia Italiana
Autore della voce: Giovanni Gallavotti
versione preliminare: giugno 1984
versione 2: gennaio 1986
versione 3: febbraio 1986

§1 *Introduzione*
§2 *La dinamica microscopica*
§3 *Medie temporali ed ipotesi ergodica*
§4 *Tempi di ricorrenza ed osservabili macroscopiche.*
§5 *"Insiemi Statistici" (o "Ensembles" o "Monodi")*
§6 *Modelli di termodinamica. Gli insiemi microcanonico e canonico e la loro relazione con l' ipotesi ergodica.*
§7 *L' avvicinamento all' equilibrio e l' equazione di Boltzmann. Ergodicità ed irreversibilità.*

§1 Introduzione

La MS pone il problema di dedurre le proprietà macroscopiche della materia dalla ipotesi atomica intesa in senso stretto.

La materia è, in questa ipotesi, costituita da aggregati di atomi o molecole che si muovono secondo le leggi della meccanica classica →, ovvero della meccanica quantistica → e la corrispondente teoria prende il nome di MS classica o MS quantistica.

Consiste, dunque, la materia, di un grandissimo numero N di particelle essenzialmente puntiformi ($N = 6.02\ 10^{23}$ particelle per grammomolecola= "numero di Avogadro"→: che implica che, ad esempio, 1 cm^3 di idrogeno o di qualsiasi altro gas perfetto in condizioni normali (1 atm a 0°C) contiene circa $2.7\ 10^{19}$ molecole) che interagiscono a mezzo di semplici leggi di forza conservative →.

Consideriamo il caso della MS classica più in dettaglio.

Uno "stato microscopico" è descritto specificando il valore degli impulsi (o, ciò che è lo stesso, delle velocità) e delle posizioni di ciascuna delle N particelle ad un dato istante: questo significa specificare $3N + 3N$ coordinate che individuano un punto dello "spazio delle fasi" → (secondo la terminologia della meccanica).

Nella visione originale di Boltzmann non sembra che le particelle venissero realmente pensate come suscettibili di assumere un continuo di stati a $6N$ dimensioni.

Piuttosto lo spazio delle fasi si pensa suddiviso in un numero finito di piccolissime cellette di uguali dimensioni, ognuna delle quali determina posizione ed impulso di ogni particella del sistema con una precisione uguale a quella "massima possibile".

Per precisione massima possibile si deve intendere la massima precisione consentita dai più perfetti apparati di misura. E si pone una questione di principio: si può o no ritenere che ogni imprecisione di misura sia perfettibile migliorando gli strumenti? se si ritiene di si allora le celle dello spazio delle fasi rappresentanti stati microscopici osservati con la massima precisione debbono essere punti e debbono essere concepiti come un continuoa $6N$ dimensioni. Però dal momento che gli atomi e le molecole non sono direttamente osservabili si possono nutrire dubbi circa la liceità dell' assunzione di perfetta misurabilità delle loro coordinate di impulso e posizione. Ad esempio nella critica ai fondamenti della meccanica classica appare il "(Principio di indeterminazione)"→ che postula come teoricamente impossibile misurare una componente p dell' impulso di una particella e la corrispondente componente q della posizione della stessa particella con precisioni rispettive δp e δq senza che

$$\delta p \delta q \geq h \tag{1.1}$$

ove $h = 6.62\ 10^{-27}$ erg sec è la "costante di Planck".

Senza addentrarsi nei problemi sollevati da queste considerazioni conviene dunque procedere immaginando che gli stati microscopici di un sistema di N particelle siano rappresentati da cellette nello spazio delle fasi costituite dai punti di coordinate:

$$\begin{cases} p_\alpha^\circ - \delta p/2 \le p_\alpha \le p_\alpha^\circ + \delta p/2 \\ p_\alpha^\circ - \delta p/2 \le q_\alpha \le q_\alpha^\circ + \delta q/2 \end{cases} \qquad \alpha = 1, \dots, 3N \tag{1.2}$$

se p_1, p_2, p_3 sono le coordinate dell' impulso della prima particella, p_4, p_5, p_6 della seconda etc., e q_1, q_2, q_3 sono le coordinate della posizione della prima particella, q_4, q_5, q_6 della seconda etc... Le coordinate p_α° e q_α° servono ad identificare il centro della generica celletta e quindi la celletta stessa.

In base alla discussione sulla taglia delle cellette si supporrà che

$$\delta p \delta q = h \tag{1.3}$$

ove h è una costante a priori arbitraria e che conviene non fissare perché può essere interessante, per i motivi appena addotti, vedere come la teoria dipenda da essa. Ha l' interpretazione di limitazione alla precisione con cui sono eseguibili misure di una coordinata di impulso e della corrispondente coordinata di posizione.

Dunque lo spazio degli stati microscopici è l' insieme delle cellette cubiche Δ di volume h^{3N} con le quali si immagina suddiviso lo spazio continuo delle fasi. Per ipotesi non ha senso porsi il problema di tentare di individuare con maggiore precisione lo stato microscopico.

La ottimistica situazione classica in cui sono possibili misure perfette di impulso e posizione simultanee sarà ottenuta considerando, nella teoria più generale, il limite per h che tende a zero.

La realtà di questa situazione in cui $h = 0$, anche volendo ignorare quanto si è appreso dalla meccanica quantistica, è non direttamente verificabile a causa della impossibilità pratica di osservare con infinita precisione (o solo con "grande" precisione) un singolo atomo.

§2 *La dinamica microscopica.*

L' ipotesi atomica, oltre a supporre l' esistenza di atomi e molecole suppone anche che tali particelle si muovano seguendo una legge di moto → deterministica.

Questa ipotesi può essere imposta pensando che sia definita una legge di trasformazione S:

$$S\Delta = \Delta' \tag{2.1}$$

che trasforma le cellette dello spazio delle fasi le une nelle altre e che descrive la dinamica del sistema: se al tempo t il sistema è individuato microscopicamente dalla celletta Δ dello spazio delle fasi, al tempo $t+\tau$ sarà individuato dalla celletta Δ'. Qui τ è una unità di tempo estremamente piccola rispetto agli intervalli di tempo macroscopici sui quali si osserva il sistema dato e, tuttavia, è un intervallo di tempo accessibile a misurazione diretta, almeno in linea di principio.

La legge di evoluzione S non è arbitraria ma deve verificare alcune proprietà fondamentali: e precisamente deve verificare le leggi della meccanica newtoniana onde realmente tradurre il principio del meccanicismo insito nell' ipotesi atomica.

Questo significa, anzitutto, che ad ogni celletta Δ dello spazio delle fasi, sono associate tre grandezze fondamentali: le energie "cinetica", "potenziale" e "totale" denotate rispettivamente $T(\Delta), \Phi(\Delta), E(\Delta)$.

Se, per semplicità, si suppone che il sistema consista di N particelle identiche di massa m e interagenti due a due via una forza conservativa di energia potenziale φ e se Δ è la celletta individuata da $(\underline{p}^\circ, \underline{q}^\circ)$, vedi (1.2), tali grandezze sono definite rispettivamente da:

$$\begin{aligned} T(\underline{p}^\circ) \equiv T(\Delta) &= \sum_{i=1}^{N} (\underline{p}_i^\circ)^2/2m \qquad \underline{p}_i^\circ = (p_{3i-2}^\circ, p_{3i-1}^\circ, p_{3i}^\circ) \\ \Phi(\underline{q}^\circ) \equiv \Phi(\Delta) &= \sum_{i<j}^{1,N} \phi(\underline{q}_i^\circ - \underline{q}_j^\circ) \qquad \underline{q}_i^\circ = (q_{3i-2}^\circ, q_{3i-1}^\circ, q_{3i}^\circ) \\ E(\underline{p}^\circ, \underline{q}^\circ) \equiv E(\Delta) &= T(\underline{p}^\circ) + \Phi(\underline{q}^\circ) \end{aligned} \tag{2.2}$$

ove $\underline{p}_i^\circ = (\underline{p}_{3i-2}^\circ, p_{3i-1}^\circ, p_{3i}^\circ)$, $\underline{q}_i^\circ = (\underline{q}_{3i-2}^\circ, q_{3i-1}^\circ, q_{3i}^\circ)$, sono il momento e la posizione della i-ma particella, $i = 1, 2, ...N$, nello stato corrispondente al centro $(\underline{p}^\circ, \underline{q}^\circ)$ di Δ.

Rimpiazzando $\underline{p}^\circ, \underline{q}^\circ$, ossia il centro di Δ, con un altro punto $(\underline{p}, \underline{q})$ di Δ si ottengono valori $T(\underline{p}), \Phi(\underline{q}), E(\underline{p}, \underline{q})$ per l' energia cinetica, potenziale e totale diversi da $T(\Delta)$,$\Phi(\Delta)$,$E(\Delta)$: però tale differenza deve essere non osservabile; altrimenti le cellette Δ non sarebbero le più piccole osservabili, come invece si è supposto.

Se ora τ è un fissato intervallo di tempo e se si considerano le soluzioni delle equazioni del moto di Hamilton:

$$\dot{\underline{q}} = \frac{\partial E}{\partial \underline{p}}(\underline{p}, \underline{q}), \qquad \dot{\underline{p}} = -\frac{\partial E}{\partial \underline{q}}(\underline{p}, \underline{q}) \tag{2.3}$$

con dati iniziali $(\underline{p}^\circ, \underline{q}^\circ)$ al tempo 0, si troverà che il punto $(\underline{p}^\circ, \underline{q}^\circ)$ evolve nel tempo τ in un punto $(\underline{p}', \underline{q}')$. Si definisce allora S in modo che $S\Delta = \Delta'$ se Δ' è la celletta che contiene $(\underline{p}', \underline{q}')$. L' evoluzione (2.3) può condurre alcune

particelle al di fuori del volume V a loro disposizione che, per semplicità, immaginiamo cubico: si deve allora pensare sempre che le (2.3) vengano completate da "condizioni al contorno" che dicono (ad esempio) che le pareti di V sono "riflettenti", ovvero "periodiche", identificando le facce opposte di V.

Si pongono però alcune questioni di principio sulla trasformazione S che conviene non ignorare, sebbene il loro approfondimento apparirà più interessante solo in seconda lettura.

Anzitutto trascureremo la possibilità che $(\underline{p}', \underline{q}')$ sia sulla frontiera di una celletta (caso in cui Δ' non è univocamente determinato, ma che può essere evitato immaginando di deformare leggermente le pareti delle cellette).

Più importante, anzi veramente essenziale, è la questione se $S\Delta_1 = S\Delta_2$ implichi $\Delta_1 = \Delta_2$: è questa una proprietà certamente vera solo nel caso delle cellette puntiformi ($h = 0$), per il teorema di unicità delle soluzioni delle equazioni differenziali, e che ha un significato intuitivo ed un interesse evidenti per la sua connessione con la questione della reversibilità del moto.

Nella discussione di questo punto gioca un ruolo fondamentale il "teorema di Liouville" per cui la trasformazione che trasforma il dato iniziale generico $(\underline{p}, \underline{q})$ nella configurazione $(\underline{p}', \underline{q}')$ è una trasformazione che conserva il volume.

Pertanto l' insieme dei dati iniziali $(\underline{p}, \underline{q})$ in Δ evolve nel tempo τ in un insieme di punti $\tilde{\Delta}$ di volume uguale a quello di Δ.

Però $\tilde{\Delta}$ pur avendo lo stesso volume di Δ non avrà la stessa forma di parallelepipedo rettangolo di dimensioni δp o δq: sarà piuttosto, per h piccolo, un parallelepipedo ottenuto deformando Δ a mezzo di una trasformazione lineare che espande in certe direzioni e contrae in altre.

Naturalmente perché la rappresentazione degli stati di sistema, come cellette dello spazio delle fasi, sia consistente occorre scegliere il tempo τ in base ai seguenti criteri.

Si supponga che h sia molto piccolo in modo che la regione $\tilde{\Delta}$ sia pensabile come ottenuta traslando Δ ed eventualmente deformandolo a mezzo di una dilatazione o contrazione lineare in alcune direzioni (contrazione e dilatazione complessive si "bilanciano" perché il volume, come si è detto, resta costante): questa condizione si realizza facilmente se h è abbastanza piccolo, perché le soluzioni delle equazioni differenziali ordinarie sono sempre pensabili, localmente, come trasformazioni lineari (vicine all' identità per tempi piccoli).

Allora:

i) È chiaro che se S dilata e contrae in alcune direzioni, anche se di poco, devono necessariamente esistere coppie di cellette $\Delta_1 \neq \Delta_2$ per cui $S\Delta_1 = S\Delta_2$ (si pensi alla trasformazione del piano che cui trasforma (x, y) in $((1+\varepsilon)^{-1}x, (1+\varepsilon)y)$, $\varepsilon > 0$ e alla sua azione sulle celle del reticolo degli

interi). Occorre dunque che τ sia così piccolo che per la "maggior parte" delle cellette Δ:

$$\tau < \theta_+(\Delta) \tag{2.4}$$

ove $\theta_+(\Delta)$ è tale che la trasformazione S (vicina all' identità per τ piccolo, appunto) produca contrazioni ed espansioni di Δ che siano trascurabili: perché solo così la $S\Delta_1 = S\Delta_2$ con $\Delta_1 \neq \Delta_2$ potrà verificarsi solo per una frazione piccola delle cellette e si potrà pensare che tali eventualità siano trascurabili.
ii) Ma τ non può essere troppo piccolo se si vuole mantenere coerentemente il punto di vista che gli stati microscopici sono descritti da cellette. Infatti alla celletta Δ è associata una scala di tempo naturale $\theta_-(\Delta)$ che può essere definita come il tempo minimo perché Δ sia distinguibile dalla celletta in cui si evolve nel tempo $\theta_-(\Delta)$: e τ deve essere certamente più grande di questa scala di tempo "minima":

$$\theta_-(\Delta) < \tau \tag{2.5}$$

Riassumendo, dunque, per poter definire la dinamica correttamente come una trasformazione che permuti fra loro le cellette occorre che h (cioè la dimensione delle cellette) sia piccola e che τ sia scelta in modo che:

$$\theta_- = \text{“}\max_{\Delta}\text{”}\, \theta_-(\Delta) \leq \tau \leq \theta_+ = \text{“}\min_{\Delta}\text{”}\, \theta_+(\Delta) \tag{2.6}$$

ove le virgolette significano che il massimo ed il minimo devono essere scelti al variare di Δ nella "maggior parte" delle cellette, nel cui insieme si possa quindi ritenere che $\Delta_1 \neq \Delta_2$ implichi $S\Delta_1 \neq S\Delta_2$.
È facile rendersi conto che se φ è un potenziale "ragionevole" (un tipico modello per φ è il modello di Lennard–Jones $\varphi(r) = 4\varepsilon((r_0/r)^{12} - (r_0/r)^6)$, ove ε è l' "intensità" di φ e r_0 ne è la "portata"), si avrà in genere:

$$\lim_{h\to 0} \max \theta_-(\Delta) = 0 \tag{2.7}$$

mentre per h piccolo il membro di destra della (2.6) (che ha carattere puramente geometrico) diviene indipendente da h.
Dunque nel limite $h \to 0$ è possibile scegliere τ in modo che valgano le (2.4),(2.5) e cioè siano soddisfatti i criteri sopra elencati come necessari per la consistenza della descrizione degli stati del sistema a mezzo di cellette.
Però se $h > 0$, e a posteriori si deve pensare che $h = 6.62\ 10^{-27}$ erg sec. la questione diviene delicata, anzi delicatissima, soprattutto perché ancora non sappiamo cosa si debba intendere per "maggior parte" delle cellette. In realtà in base ai risultati della teoria è possibile valutare l' effetto sui risultati stessi della presenza di coppie di cellette con $\Delta_1 \neq \Delta_2$ per cui

$S\Delta_1 = S\Delta_2$. A questo punto si dovrebbe dunque, logicamente, posporre il problema fino a che le conseguenze delle ipotesi che andiamo ponendo non permettano di riesaminare la questione.

Conviene però, per dare un'idea degli ordini di grandezza coinvolti nel problema e della sua delicatezza, anticipare alcuni dei risultati fondamentali e dare una stima di θ_- e θ_+: il lettore che preferirà immaginare che $h = 0$ in base ad una interpretazione dogmatica della ipotesi atomica potrà evitare la discussione che segue e procedere nell' analisi dei fondamenti della MS considerando il limite per $h \to 0$ della teoria che segue.

Tuttavia, è bene sottolinearlo, questa è una semplificazione illusoria e che evita un problema che oggi sappiamo essere tutt'altro che vano: il presupporre che si accetti per evidente un fatto che tale non è da un punto di vista sperimentale, e cioè che abbia senso almeno in linea di principio misurare con esattezza posizione e velocità di un enorme numero di molecole (o anche di una sola). E questa difficoltà, è utile ricordare, fu proprio quella per cui nel secolo scorso a molti risultò impossibile accettare l' ipotesi atomica.

Tornando al problema di fornire un'idea degli ordini di grandezza di θ_- e θ_+ si può pensare di interpretare "max" e "min" in (2.6) come valutati considerando le cellette Δ "tipiche" nelle quali l' impulso p e la velocità v della grande maggioranza delle particelle assumano i loro "valori medi": vedremo che la teoria degli insiemi statistici condurrà ad una naturale distribuzione di probabilità sulle cellette, in base alla quale sarà possibile calcolare i valori medi di varie grandezze, in una situazione stazionaria, in termini di quantità quali la temperatura assoluta T, la massa m delle particelle, il numero di particelle N, il volume V a disposizione del sistema.

La relazione fondamentale da usare è quella che fornirà il valore medio dell' impulso e della velocità $\overline{p}$ e $\overline{v}$:

$$\overline{p} = m\overline{v} = \sqrt{3mkT}, m\overline{v}^2 = 3kT \tag{2.7}$$

ove $k = 1.38\ 10^{-16}$ erg/$^\circ K$ è la costante di Boltzmann e T è la temperatura assoluta ed m è la massa delle particelle.

Altre grandezze rilevanti sono i parametri caratteristici dell' interazione, cioè ε, la "intensità" con dimensioni di un'energia, e r_0, la "portata" con dimensioni di una lunghezza: dalla applicazione della teoria generale segue, indipendentemente dalla forma di $\varphi(r)$ (purché ragionevole; ad esempio si può considerare il modello di Lennard–Jones su menzionato), che $\varepsilon \cong kT_c^o$ ove T_c^o è la temperatura critica di liquefazione e r_0 è di qualche diametro molecolare (dai $2\ 10^{-8}$ cm ai $4\ 10^{-8}$ cm nelle molecole più semplici quali H_2, He, O_2, CO_2 etc).

Come prima quantità stimiamo $\theta_+(\Delta)$ in una celletta tipica in cui si possa assumere che le particelle evolvono nel tempo senza che alcuna di esse subisca collisioni multiple.

In questa situazione la variazione relativa di una dimensione lineare di Δ nel tempo τ sarà, per τ piccolo, proporzionale a τ e potrà dipendere da $\varepsilon, m, r_0, \overline{v}$: i numeri puri proporzionali a τ, e legati alle dilatazioni, che si possono formare con tali grandezze sono $\tau(\varepsilon/mr_0^2)^{1/2}$ e $\tau(m\bar{v}^2/mr_0^2)^{1/2}$. Pertanto le dilatazioni saranno trascurabili se, ricordando che $m\overline{v}^2 = 3kT$ e posto $\varepsilon \equiv kT_c$:

$$\tau < \min\left(\left(\frac{mr_0^2}{kT_c}\right)^{1/2}, \left(\frac{mr_0^2}{kT}\right)^{1/2}\right) \equiv \theta_+ \tag{2.8}$$

e la condizione $\tau^2\varepsilon/mr_0^2 < 1$ significa che durante una collisione non c'è espansione, mentre la seconda, $\tau < r_0/\bar{v}$, esprime che il tempo τ è piccolo rispetto alla durata delle collisioni.

La stima di $\theta_-(\Delta)$ può essere basata invece sulle seguenti considerazioni. Data Δ, le coordinate p_α e q_α dei punti della celletta Δ variano nel tempo τ, in base alle equazioni del moto, di:

$$\begin{aligned} |\delta q_\alpha| &\cong |\tau \frac{\partial E}{\partial p_\alpha}(\underline{p}, \underline{q})| = \tau \frac{\delta_\alpha^{(1)} E}{\delta p} \\ |\delta p_\alpha| &\cong |-\tau \frac{\partial E}{\partial q_\alpha}(\underline{p}, \underline{q})| = \tau \frac{\delta_\alpha^{(2)} E}{\delta q} \end{aligned} \tag{2.9}$$

ove $\delta_\alpha^{(1)}E$, $\delta_\alpha^{(2)}E$ sono le variazioni della energia E nella celletta Δ se le coordinate p_α o q_α variano di δp o δq, cioè delle dimensioni lineari di Δ, le altre restando fisse.

Definendo allora l' "indeterminazione dell' energia", denotata $\delta E(\Delta)$, nella celletta Δ come:

$$\delta E(\Delta) = \max_\alpha(\delta_\alpha^{(1)}E, \delta_\alpha^{(2)}E) \tag{2.10}$$

Si vede che la quantità di tempo minima $\theta_-(\Delta)$ che occorre attendere perché la celletta evolva in una celletta distinguibile da Δ stessa è quella per cui:

$$\theta_-(\Delta)\max_\alpha \tfrac{\delta_\alpha^{(1)}E}{\delta p} \geq \delta q, \quad \text{ovvero} \quad \theta_-(\Delta)\max_\alpha \tfrac{\delta_\alpha^{(2)}E}{\delta q} \geq \delta p \tag{2.11}$$

perché appunto δp e δq sono le dimensioni lineari di Δ e (2.11) dice che almeno uno dei lati di Δ si è mosso di una quantità pari alla sua lunghezza.

Ricordando che $\delta p\, \delta q = h$ si ha, dalle (2.10),(2.11):

$$\theta_-(\Delta)\delta E(\Delta) \geq h \tag{2.12}$$

per cui si potrà scegliere $\theta_- \equiv \delta t$ in modo che se:

$$\delta E = \text{“}\min_\Delta\text{”}\ \delta E(\Delta) \tag{2.13}$$

sia:

$$\delta t \delta E = h \tag{2.14}$$

Volendo ora trovare in base alle (2.8),(2.14) un intervallo (θ_-, θ_+) ammissibile per τ si può immaginare che $\delta p \simeq \bar{p}$ e quindi $\delta E = \overline{p}\,\delta p/m \cong \bar{p}^2/m = 3kT$ e dunque:

$$\theta_- \equiv h/kT \tag{2.15}$$

che fornisce anche una assai interessante interpretazione alla scala di tempo h/kT come il tempo necessario perché una celletta, tipica fra quelle che descrivono stati microscopici a temperatura T, divenga distinguibile da se stessa.

In altri termini θ_+ è determinato dalla grandezza delle $\dot{p}, \dot{q}$, cioè dalle derivate prime della hamiltoniana mentre θ_- è legato all' espansione nello spazio delle fasi e cioè alle derivate seconde della hamiltoniana.

Con un po' di algebra si trova, da (2.8), (2.15):

$$\theta_+/\theta_- = (mr_0^2 kT_c/h^2)^{1/2} \min(T/T_c, (T/T_c)^{1/2}).$$

Pertanto è chiaro che la condizione $\theta_+/\theta_- > 1$, necessaria perché esista τ verificante (2.6), e cioè perché sia consistente la descrizione degli stati microscopici in termini di cellette, sarà verificata per T grande, diciamo $T \gg T_0$ ma non per T piccolo. E dalla espressione appena dedotta per θ_+/θ_- si trova:

$$T_0/T_c = \max\Big(h(mr_0^2 kT_c)^{-1/2}\,,\, \big(h(mr_0^2 kT_c)^{-1/2}\big)^2\Big) \tag{2.16}$$

La seguente tabella dà un'idea degli ordini di grandezza ed è elaborata avendo scelto $h = 6.62\ 10^{-27}$ erg/sec (un risultato non molto diverso per h e quindi $\delta\tau$, si otterrebbe se, ignorando la costante di Planck suggerita dalla meccanica quantistica, si scegliesse $\delta p \ll \overline{p}$, $\delta q \ll \sqrt[3]{V/N}$. Infatti così la $\delta p\,\delta q \ll \overline{p}\sqrt[3]{V/N}$ risulterebbe in casi ragionevoli, (1 cm^3 di idrogeno, $m = 3.34\,10^{-24}$ g, $T = 273°\mathrm{K}, N = 2.7\,10^{19}$, $k = 1.38\,10^{-16}$ erg/°K), dello stesso ordine di grandezza della costante di Planck: $\delta p\,\delta q \ll 2.04\ 10^{-25}$ erg sec. L' ordine di grandezza corrispondente di θ_- è $\theta_- \cong 5.43\ 10^{-12}$ sec.

La vicinanza fra le stime di θ_- e θ_+ per $T \cong T_c$, che si desume dalla tabella, mostra che la questione della coerenza della rappresentazione in termini di cellette, presa in senso stretto, è dipendente in modo molto delicato da h e in ogni caso se $h \neq 0$ è destinata a non essere consistente se $T \to 0$ (perché $\theta_- \to \infty$) e $\varepsilon \neq 0$ ($\theta_+ \to \sqrt{mr_0^2/\varepsilon}\ < \infty$).

	r_0	ε	T_0/T_c	T_c^{vera}	A	B	m
	10^{-8}cm	10^{-14}erg		°K	10^{-11} cgs	cgs	1.17 10^{-24}g
H_2	2.76	0.57	4.3	33.2	2.46	26.7	2
CO_2	3.23	5.25	0.12	304.1	36.	42.7	44
He	2.65	0.08	15.	5.19	0.33	23.5	4
N_2	3.12	2.17	0.25	126.	13.4	38.6	28
O_2	2.93	2.65	0.23	154.3	13.6	31.9	32

le colonne A, B riportano i dati empirici, direttamente accessibili agli esperimenti, della equazione di stato di Van der Waals ($n = N/$(numero di Avogadro)$= N/N_A$ =numero di moli, $R = kN_A$), si veda Viriale per le (*) e (**) successive:

$$(P + An^2/V^2)(V - nB) = nRT \qquad (*)$$

che si assume per dedurre ε e r_0 via le relazioni

$$(B/N_A) = 4\frac{4\pi}{3}\left(\frac{r_0}{2}\right)^3 = 4v_0 \qquad A/N_A^2 = \frac{32}{3}\varepsilon v_0 \qquad (**)$$

che conducono alle espressioni $r_0 = (3B/2\pi N_A)^{1/3}, \varepsilon = 3A/8BN_A) = 81\,kT_c/64$.

§3 Medie temporali ed ipotesi ergodica.

Si è dunque condotti a pensare che la descrizione di un sistema meccanico di N particelle identiche di massa m sia esprimibile (almeno a temperature non troppo basse, $T > T_0$ cfr. (2.16)) in termini di una "funzione energia" definita sullo spazio delle fasi a $6N$ dimensioni e di una suddivisione dello spazio delle fasi in cellette Δ di uguale volume h^{3N} legato alla precisione con cui si pensa di poter misurare impulsi e posizioni o intervalli di tempo ed energie.

L' evoluzione temporale osservata su intervalli di tempo multipli di una unità τ grande rispetto alla scala di tempo $\delta\tau$ associata alla decomposizione in cellette dalla relazione (2.14) e piccola rispetto alla scala di tempo di

"collisione", (2.8), sarà pensata come una permutazione delle cellette di data energia: si trascura qui il fatto, discusso al §2, che possano esistere cellette diverse che evolvono nella stessa celletta.

Date le premesse ed ipotesi sopra enunciate ci si domanda quale sia il comportamento qualitativo di un sistema con energia "fissata" macroscopicamente, e cioè compresa fra $E - DE$ ed E, osservato ad intervalli di tempo τ (con DE quantità macroscopicamente piccola ma tale che $DE \gg \delta E = h/\delta\tau$, cfr. (2.14), (2.15)).

Boltzmann suppose che, nei casi interessanti, valesse l' "ipotesi ergodica" $\rightarrow$ secondo la quale l' azione di S è la più semplice possibile: ossia S è una *permutazione ad un solo ciclo* delle $\mathcal{N}$ cellette della data energia:

$$S\Delta_k = \Delta_{k+1} \qquad k = 1, 2, \ldots, \mathcal{N} \tag{3.1}$$

se le cellette sono numerate opportunamente (e $\Delta_{\mathcal{N}+1} \equiv \Delta_1$).

In altre parole, all' evolversi del tempo, ogni celletta evolve visitando successivamente tutte le altre di uguale energia.

Il fondamento di questa celebre e contestata ipotesi è la sua semplicità concettuale: dice che nel sistema in questione tutte le cellette della medesima energia sono equivalenti.

Ci sono casi in cui questa ipotesi è manifestamente falsa: se ad esempio il sistema è racchiuso in un contenitore sferico perfetto l' evoluzione conserva il "momento angolare" $\rightarrow$ ($\underline{M}(\Delta) = \sum_{i=1}^{N} \underline{p}_i^\circ \wedge \underline{q}_i^\circ$), per cui cellette con diverso momento angolare non possono evolvere l' una nell' altra. Questo è, sostanzialmente, il caso più generale in cui la ipotesi ergodica non è corretta: se l' evoluzione non è una permutazione ad un solo ciclo la si può decomporre in cicli e corrispondentemente definire sulle cellette una funzione A che ad ogni celletta associa un valore uguale per tutte le cellette di uno stesso ciclo e diverso fra cellette di cicli diversi. Evidentemente la funzione A così definita è una costante del moto che gioca lo stesso ruolo del momento angolare dell' esempio precedente.

Dunque, se l' ipotesi ergodica non fosse valida, nel sistema varrebbero altre leggi di conservazione oltre alla legge di conservazione dell' energia: in questi casi sarebbe naturale immaginare di fissare le quantità conservate e domandarsi quali siano le proprietà qualitative dei moti di energia E quando tutte le altre costanti del moto siano anche fissate: ovviamente, ora, il moto sarà per definizione una semplice permutazione ciclica di tutte le cellette permesse dai valori prefissati dell' energia e delle altre costanti del moto. Quindi, in un certo senso, l' ipotesi ergodica non sarebbe restrittiva.

Questa osservazione, come ben si avvide Boltzmann stesso, non sminuisce affatto il problema relativo all' ipotesi ergodica e, al contrario, ne mette in luce alcuni aspetti sottili e profondi.

Infatti la decomposizione di S in cicli ("decomposizione ergodica di S" $\rightarrow$) potrebbe essere così complicata da rendere praticamente impossibile la costruzione e, cioè, inutile: questo avverrebbe se le regioni dello spazio delle fasi corrispondenti ai vari cicli fossero, in qualche direzione almeno, di taglia microscopica ovvero di taglia molto minore di quella macroscopica ovvero fossero molto irregolari su scala microscopica (a differenza di quanto accade nell' esempio semplice, prima dato, della conservazione del momento angolare).

Non è invero inconcepibile che possano esistere in casi interessanti complicatissimi integrali primi privi di interpretazione fisica macroscopica diretta.

Dunque il "problema ergodico" $\rightarrow$, cioè la verifica dell' ipotesi ergodica, in casi in cui non ci siano particolari ragioni di simmetria che implichino l' esistenza di ovvii integrali primi è un problema che resta da studiare caso per caso. Una sua soluzione soddisfacente sarebbe la dimostrazione della sua validità in senso stretto ovvero della possibilità di individuare i cicli di S mediante superfici di livello di funzioni semplici (ossia di integrali primi semplici) dotate di significato fisico macroscopico (ad esempio in termini di leggi di conservazione, come nel caso già illustrato del momento angolare).

È bene sottolineare che non si deve pensare che non esistano altri esempi semplici ed interessanti in cui l' ipotesi ergodica è manifestamente falsa. L' esempio più classico è la catena chiusa di oscillatori armonici descritti da:

$$T = \sum_{i=1}^{N} p_i^2/2m \quad \Phi = \sum_{i=1}^{N} m(q_{i+1} - q_i)^2/2 \tag{3.2}$$

in cui, per semplicità, $q_{N+1} = q_1$ (condizione di chiusura o "periodicità").

In questo caso esistono N integrali primi:

$$A_k = (\underline{p} \cdot \underline{\eta}_k)^2 + \omega(k)^2(\underline{q} \cdot \underline{\eta}_k)^2 \quad k = 1, 2, \ldots, N \tag{3.4}$$

ove $\underline{\eta}_1$, $\underline{\eta}_2, \ldots, \underline{\eta}_N$ sono N vettori ortonormali opportuni ("modi normali" $\rightarrow$) e $\omega(k)$ sono le "pulsazioni proprie" $\rightarrow$ della catena:

$$\omega(k)^2 = 2(1 - \cos 2\pi k/N) \tag{3.5}$$

Allora cellette Δ e Δ' dello spazio delle fasi per cui i vettori $\underline{A}(\Delta) = (A_1(\Delta), A_2(\Delta), \ldots, A_N(\Delta))$ e $\underline{A}(\Delta')$ non coincidono non possono appartenere allo stesso ciclo e dunque il sistema non è ergodico.

Tuttavia Boltzmann ritenne che circostanze come l' ultima descritta si dovessero considerare eccezionali: e converrà non approfondire subito il problema ergodico, sia per la sua difficoltà sia per vedere come si possa procedere oltre nella formulazione della MS classica.

Dato dunque un sistema meccanico di N particelle identiche (per semplicità) si consideri il problema dello studio di una data "grandezza osservabile." $f(\underline{p},\underline{q})$ definita sullo spazio delle fasi.

La prima quantità importante che si può studiare, e spesso l' unica necessaria, è il valore medio di f:

$$\overline{f(\Delta)} = \lim_{T\to\infty} \frac{1}{T} \sum_{k=0}^{T} f(S^k\Delta) \tag{3.6}$$

ove $f(\Delta) = f(\underline{p},\underline{q})$ se $(\underline{p},\underline{q})$ è un punto che individua Δ.

Se $\Delta_1 = \Delta, \Delta_2, \ldots, \Delta_{\mathcal{N}}$ è il ciclo cui appartiene la celletta Δ chiaro che:

$$\overline{f(\Delta)} = \frac{1}{\mathcal{N}} \sum_{k=1}^{\mathcal{N}} f(\Delta_k) \tag{3.7}$$

e nel caso ergodico il ciclo consiste nell' insieme di tutte le cellette di energia uguale a quella di Δ.

Se l' energia del sistema è determinata a meno di un errore macroscopico DE trascurabile da un punto di vista macroscopico (ma grande rispetto δE), le cellette di energia fra $E - DE$ ed E saranno suddivise in cicli di energia variabile ma su ciascuno di essi la f dovrebbe avere lo stesso valore medio (a meno di variazioni trascurabili). Quindi, se J denota il dominio delle variabili $(\underline{p},\underline{q})$ in cui vale la $(E - DE \le E(\underline{p},\underline{q}) \le E)$, si dovrà avere:

$$\overline{f(\Delta)} = \int_J f(\underline{p},\underline{q})\ \mathrm{d}p\ \mathrm{d}q / \int_J d\underline{p}\, d\underline{q} \tag{3.8}$$

Ricordando, infatti, che le cellette hanno tutte uguale volume la (3.8) segue immediatamente dalla (3.7) e dalla indipendenza, supposta sopra, di $\overline{f(\Delta)}$ da $E(\Delta)$, **se** h è così piccolo che la somma sulle cellette può essere sostituita dagli integrali.

Questa relazione, che Boltzmann congetturò essere vera "salvo casi eccezionali" (quali la catena di oscillatori perfetta prima descritta) e scrisse nella forma suggestiva:

$$\lim_{T\to\infty} \frac{dt}{T} = \left.\frac{d\underline{p}d\underline{q}}{\int d\underline{p}d\underline{q}}\right|_E \tag{3.9}$$

si legge "la media temporale di un' osservabile è uguale alla sua media sulla superficie di energia costante". Come vedremo, cfr. §6, la (3.9) fornisce la base euristica del "modello microcanonico" della termodinamica classica.

Si noti che se è vera la (3.9), cioè se è vera la (3.8), il valore medio di un' osservabile dipende solo da E e non dalla particolare celletta Δ in cui il sistema si trova inizialmente.

Quest' ultima proprietà è certamente un prerequisito per ogni programma che intenda dedurre le proprietà macroscopiche della materia dall' ipotesi atomica, in quanto è ovvio che tali proprietà non possono dipendere dai dettagli delle proprietà microscopiche della configurazione Δ in cui il sistema si viene a trovare ad un dato istante.

Infine è opportuno notare che nella (3.7) non appare più la dinamica microscopica: questa infatti è implicita nella numerazione delle cellette fatta in modo che $\Delta_1, \Delta_2, \Delta_3, \ldots$ siano le cellette in cui Δ evolve successivamente ad intervalli di tempo τ. Però nella (3.7) è chiaro che l' ordine di tale numerazione non è importante e lo stesso risultato si otterrebbe se le cellette di data energia venissero numerate diversamente.

Si vede così il fascino dell' ipotesi ergodica che sembra liberarci dalla necessità di conoscere i dettagli della dinamica microscopica ai fini del calcolo delle medie delle osservabili. La illusorietà di questa concezione, chiara già a Boltzmann, è messa in luce nei successivi §4,6,7.

§4 *Tempi di ricorrenza ed osservabili macroscopiche.*

Nelle applicazioni ha grande importanza sapere valutare la velocità con cui il limite f viene raggiunto: perché la (3.8) sia utile occorre che il limite in (3.6) sia raggiunto in un tempo t, pur lungo rispetto a τ, ma molto corto rispetto ai tempi rilevanti per le osservazioni macroscopiche che si vogliono eseguire sul sistema. Infatti è solo su scale di tempo dell' ordine di t o più lunghe che l' osservabile f appare costante e uguale al suo valore medio.

È perfettamente concepibile una situazione in cui il sistema è ergodico ma il valore $f(S^k\Delta)$ oscilla talmente, lungo la traiettoria, che il valore medio di f viene raggiunto su scale di tempo dell' ordine di grandezza del tempo necessario a visitare tutta la superficie di energia costante che è necessariamente enorme.

Ad esempio riferendosi ai calcoli di ordine di grandezza discussi alla fine del §2, c.f.r. i valori di δp, δE prima della (2.15) e la (2.15) stessa, bisogna calcolare il numero di cellette di volume h^{3N} contenute nella regione fra E ed $E + \delta E$.

Se la superficie della sfera unitaria a d dimensioni si scrive $2\sqrt{\pi}^d\Gamma(d/2)^{-1}$ (essendo Γ la "funzione gamma di Eulero"), il volume della regione menzionata, nel caso in cui h sia molto piccolo, può essere calcolato usando coordinate polari nello spazio degli impulsi p. Le cellette in questione sono quelle tali che $P \equiv \sqrt{\Sigma_i \underline{P}_i^2}$ varia da $P = \sqrt{2mE}$ a $P+\delta P = \sqrt{2m(E+\delta E)}$.

Se si introducono, c.f.r. §2, (2.15) etc., le quantità:

$$\begin{aligned} P =&\sqrt{2mE} \quad \delta p = \overline{p} = \sqrt{3mkT} \\ \delta E =&3kT = \overline{p}\delta p/m \quad \frac{E}{N} = \frac{3kT}{2} \\ \delta P =&\frac{\overline{p}\delta p}{P} = \frac{\delta p}{\sqrt{N}} \quad \delta q = \left(\frac{V}{N}\right)^{1/3} \end{aligned} \tag{4.1}$$

ove k è la costante di Boltzmann, $k = 1.38\ 10^{-16}$ erg/°K, T la temperatura assoluta, V il volume occupato dal gas ed N il numero di particelle, si trova che il volume in questione è, se si pone $h = \delta p \delta q$ e si usa la formula diStirling per valutare $\Gamma(3N/2)$:

$$\begin{aligned} w &= V^N \sqrt{2mE}^{3N-1}\, \delta P\, \sqrt{\pi}^{3N}/2\Gamma(3N/2) = \\ &= (N\delta q^3)^N (\sqrt{N}\delta p)^{3N-1} \delta P \sqrt{\pi}^{3N}/2\Gamma(3N/2) = \\ &= (\delta p \delta q)^{3N} N^{\frac{5}{2}N-1} \sqrt{\pi}^{3N}/2\Gamma(3N/2) \cong h^{3N} \frac{N^{\frac{5}{2}N-1}}{N^{\frac{3}{2}N-\frac{1}{2}}} (2\pi e/3)^{\frac{3N}{2}} \frac{1}{2\sqrt{\pi}\left(\frac{3}{2}\right)^{3/2}} \end{aligned} \tag{4.2}$$

Quindi il numero $\mathcal{N}$ di cellette è w/h^{3N} ed il tempo di ricorrenza), se il sistema si muovesse ergodicamente sulla superficie di energia E:

$$T_{\text{ricorrenza}} = \mathcal{N}\tau \cong \mathcal{N}\frac{h}{kT} \cong N^{N-1/2}\frac{h}{kT}\left(\frac{2\pi e}{3}\right)^{3N/2} \tag{4.3}$$

Come discusso al §2 l' ordine di grandezza di h/kT è, se $T = 300$°K, di circa 10^{-14} sec (per i nostri scopi non farà alcuna differenza che per h si usi la formula $h = \delta p \delta q$ con δp e δq dati in (4.1) con $V = 1$ cm^3, $N = 2.7\ 10^{19}$, $m = 3.34\ 10^{-24}$g = massa della molecola di idrogeno, ovvero che si usi il valore della costante di Planck).

Dunque il tempo di ricorrenza in (4.3) è inimmaginabilmente più lungo della età dell' universo non appena N supera il valore di qualche centinaio (anche se molto minore del numero diAvogadro) e se T viene scelto = 0°C: per 1 cm^3 di idrogeno a 0°C si ha $N \simeq 10^{19}$ e $T_{\text{ricorrenza}} = 10^{-14} \cdot 10^{10^{19}}$ sec, mentre l' età dell' universo è (solo) $\sim 10^{17}$ sec!

L' idea di Boltzmann per conciliare l' ergodicità con la rapidità dell' avvicinamento all' equilibrio è che le osservabili interessanti, "osservabili macroscopiche", siano tali da avere valore praticamente costante sullo spazio delle fasi di data energia eccezion fatta per una piccolissima frazione ε delle cellette: pertanto il tempo necessario affinché il valore medio asintotico venga raggiunto sarà non già dell' ordine di grandezza dell' iperastronomico tempo di ricorrenza microscopica bensì dell' ordine di

$T' = \varepsilon T_{\text{ricorrenza}}$. E si deve pensare che $\varepsilon \to 0$ al crescere del numero di particelle e che T' sia di moltissimi ordini di grandezza più piccolo del tempo di ricorrenza in modo da essere macroscopicamente osservabile! Si veda §7.

Esempi di osservabili macroscopiche interessanti sono:

1) il rapporto tra il numero di particelle, che si trovano in un cubetto Q, diviso per il volume di Q: questa grandezza sarà denotata $\rho(Q)$ e il suo valore medio ha l' interpretazione di "densità" in Q;

2) la somma delle energie cinetiche delle particelle $T(\Delta) = \Sigma_i \underline{p}_i^2/2m$;

3) la energia potenziale totale del sistema: $\Phi(q) = \Sigma_{i<j}\varphi(\underline{q}_i - \underline{q}_j)$;

4) il numero di particelle che si trovano in un cubetto Q aderente alla parete del contenitore V del sistema e che hanno una componente della velocità negativa lungo la normale interna e compresa fra $-v$ e $-(v + \mathrm{d}v)$, con $v > 0$. Tale numero diviso per il volume di Q è la "densità $n(Q, v)\,\mathrm{d}v$ delle particelle di velocità normale v che stanno per collidere con la parete esterna del cubetto Q". La somma, sui valori di v e sui cubetti Q adiacenti alla frontiera del contenitore V delle quantità che appaiono nella seguente formula:

$$\sum_Q \int_{v>0} \mathrm{d}v n(Q, v)(2mv)(vs/S) = P(\Delta) \tag{4.4}$$

con s = superficie di una faccia del cubetto Q, e S = superficie del contenitore, è l' impulso trasferito dal gas alla parete per unità di tempo e di superficie. Infatti una molecola che urta la parete cambia la componente normale dell' impulso di $2mv$ (da mv a $-mv$) e il numero di collisioni per unità di tempo dovute a particelle di velocità normale v è $n(Q, v)vs\,\mathrm{d}v$. La quantità (4.4) è un'osservabile che sarà denotata $P(\Delta)$: il suo valore medio la forza media esercitata dal gas sulla parete per unità di superficie, ossia ha l' interpretazione di "pressione".

5) Il prodotto $\rho(Q) \cdot \rho(Q')$, il cui valore medio è la "funzione di correlazione" fra il cubetto Q ed il cubetto Q': fornisce informazioni sulla probabilità congiunta di trovare una particella in Q e, simultaneamente, una in Q'.

§5 "Insiemi statistici" (o "Ensembles" o "Monodi") e modelli di termodinamica.

Da un punto di vista più generale, senza supporre la validità della ipotesi ergodica, è chiaro che il valore medio di una osservabile esisterà sempre e sarà uguale al valore medio sul ciclo al quale appartiene il dato iniziale, c.f.r. (3.7).

Per descrivere più quantitativamente questa osservazione si introduce la

nozione di "misura stazionaria": è una funzione che ad ogni celletta associa un numero $\mu(\Delta)$ ("misura di Δ") in modo che:

$$\mu(\Delta) \geq 0 \qquad \sum_{\Delta} \mu(\Delta) = 1 \qquad \mu(\Delta) = \mu(S\Delta) \tag{5.1}$$

Si suol dire che μ è una "misura di probabilità invariante" o "stazionaria" sullo spazio delle fasi (o, meglio, sulle cellette dello spazio delle fasi).

Se si immagina di ricoprire lo spazio delle fasi con un fluido in modo che la massa di fluido in Δ sia $\mu(\Delta)$ e se i punti dello spazio delle fasi vengono fatti muovere secondo la legge S della dinamica, allora il fluido appare immobile, ossia la sua distribuzione sullo spazio delle fasi resta invariante (o stazionaria) al variare del tempo: questo motiva il nome usato per μ.

E' chiaro che $\mu(\Delta)$ deve avere lo stesso valore su tutte le cellette di uno stesso ciclo $\mathcal{C}_\alpha$ di S (qui α è un indice che distingue i vari cicli di S). Se $\mathcal{N}(\mathcal{C}_\alpha)$ è il numero di cellette del ciclo α si deve avere quindi $\mu(\Delta) = p_\alpha/\mathcal{N}(\mathcal{C}_\alpha)$, con $p_\alpha \geq 0$ e tale che $\sum_\alpha p_\alpha = 1$, per $\Delta \in \mathcal{C}_\alpha$.

Conviene definire per ogni ciclo $\mathcal{C}_\alpha$ di S la misura stazionaria μ_α ponendo:

$$\begin{cases} \mu_\alpha(\Delta) = 1/\mathcal{N}(\mathcal{C}_\alpha) & \text{se } \Delta \in \mathcal{C}_\alpha \\ = 0 & \text{altrimenti} \end{cases} \tag{5.2}$$

il che permette di pensare ogni misura invariante μ come combinazione lineare delle misure μ_α associate ai vari cicli di S:

$$\mu(\Delta) = \sum_\alpha p_\alpha \mu_\alpha(\Delta) \tag{5.3}$$

Questa decomposizione della più generale misura invariante in somma di "misure ergodiche" (perché tali sono chiamate le misure in (5.2), in quanto concentrate su un solo ciclo di S) viene naturalmente chiamata la "decomposizione ergodica" della misura invariante μ.

Boltzmann, in un profondo lavoro del 1884, formulò l' ipotesi che le misure stazionarie m fossero interpretabili come stati di equilibrio microscopico e che, quindi, l' insieme degli stati di equilibrio macroscopico potesse essere identificato con un insieme (il termine usato da Gibbs e di uso corrente è "ensemble", mentre il termine usato da Boltzmann e ormai desueto è "monodo") di misure stazionarie definite sulle dello spazio delle fasi: un tale insieme $\mathcal{E}$ verrà qui chiamato "insieme statistico".

La identificazione fra una misura stazionaria μ sullo spazio delle fasi ed uno stato di equilibrio macroscopico avverrebbe identificando $\mu(\Delta)$ con la probabilità di trovare il sistema nella celletta Δ qualora si esegua, in un istante scelto a caso, una misura del suo stato microscopico.

Pertanto il valore medio nel tempo, nello stato di equilibrio descritto da μ, di una qualsiasi osservabile f sarebbe:

$$\overline{f} = \sum_{\Delta} \mu(\Delta) f(\Delta) \tag{5.4}$$

che è una relazione vera, per definizione, se μ è ergodica (c.f.r. righe seguenti la (5.3)) ma in generale non è direttamente collegabile alla ergodicità come già si è intravisto al §3 quando si è accennato alla lunghezza dei tempi di ricorrenza e quindi alle ulteriori ipotesi necessarie alla deduzione delle (3.8), (3.9), oltre all' ipotesi ergodica). Ritorneremo sulla (5.4) e sull' ipotesi ergodica, con la quale non deve essere confusa, al §6.

Boltzmann, nel lavoro menzionato, pose la seguente questione.

A prescindere dalla ipotesi ergodica o da qualunque tentativo di giustificazione dinamica della (5.4), si considerino tutti i possibili insiemi statistici $\mathcal{E}$ di misure stazionarie sullo spazio delle fasi.

Per ogni fissato $\mathcal{E}$ e per ogni $\mu \in \mathcal{E}$ si definisca:

$$\begin{aligned}
\Phi(\mu) = \sum_{\Delta} \mu(\Delta)\Phi(\Delta) = &\qquad \text{“energia potenziale media”} \\
T(\mu) = \sum_{\Delta} \mu(\Delta)T(\Delta) = &\qquad \text{“energia cinetica media”} \\
U(\mu) = \Phi(\mu) + T(\mu) = &\qquad \text{“energia totale media”} \\
P(\mu) = \sum_{\Delta} \mu(\Delta)P(\Delta) = &\qquad \text{“pressione”, c.f.r. (4.4)} \\
\rho(\mu) = N/V = \rho \equiv 1/v = &\qquad \text{“densità”} \\
V = \int \mathrm{d}q = &\qquad \text{“volume”}
\end{aligned} \tag{5.5}$$

ove V è il volume a disposizione del sistema, N il numero di particelle.

Ci si domanda quali insiemi statistici $\mathcal{E}$ (ossia quali "monodi" o "ensembles") abbiano la proprietà che al variare di μ in $\mathcal{E}$ le variazioni infinitesime $\mathrm{d}U$, $\mathrm{d}V$ siano legate fra loro dalla relazione (che coinvolge anche la pressione $P = P(\mu)$ e la energia cinetica media per particella $\overline{T} = T(\mu)/N$):

$$\frac{\mathrm{d}U + P\,\mathrm{d}V}{\overline{T}} = \text{ differenziale esatto} \tag{5.6}$$

almeno “nel limite termodinamico” in cui il volume $V \to \infty$ (supponendo per semplicità che il contenitore mantenga la forma cubica) e anche $N \to \infty$ in modo che la densità N/V resti fissa.

Tali insiemi statistici vennero chiamati da Boltzmann “ortodi”: sono gli insiemi statistici $\mathcal{E}$ sui quali è possibile interpretare l' energia cinetica per particella come proporzionale alla temperatura assoluta T (via una costante di proporzionalità da determinare empiricamente e denotata convenzionalmente $2/3k$, ossia: $T = 2T(\mu)/3kN$), e inoltre è possibile definire a mezzo della (5.6) una funzione $S(\mu)$, $\mu \in \mathcal{E}$, in modo che le grandezze U, ρ, T, V, P, S verifichino le relazioni che intercorrono fra le omonime grandezze della termodinamica classica, almeno nel “limite termodinamico”: in questa

identificazione la funzione S si dovrebbe identificare con l' entropia, naturalmente.

In altre parole Boltzmann pose la questione di quando e come sia possibile interpretare gli elementi di un insieme statistico stazionario $\mathcal{E}$ come stati macroscopici di equilibrio di un sistema meccanico governato dalle leggi della termodinamica classica.

Si noti che l' ipotesi ergodica combinata con le altre ipotesi usate al §3 per dedurre la (3.8), (3.9) conduce a pensare che l' insieme statistico $\mathcal{E}$ consistente delle misure μ sullo spazio delle fasi definite da:

$$\begin{aligned} &\mu(\Delta) = 1/\mathcal{N}(U,V) \qquad \text{se} \quad E(\Delta) \in (U - DE, U) \\ &\mu(\Delta) = 0 \qquad\qquad\qquad\quad \text{altrimenti} \end{aligned} \tag{5.7}$$

ove U e V sono due parametri prefissati corrispondenti alla energia totale ed al volume a disposizione del sistema, sia necessariamente un insieme statistico atto a descrivere gli stati di equilibrio macroscopico. Qui $\mathcal{N}(U,V)$ è una costante di normalizzazione da identificare come proporzionale all' integrale $\int \mathrm{d}\underline{p}\,\mathrm{d}\underline{q}$, nella (3.8), esteso alla regione di $\underline{p}$, $\underline{q}$ in cui $E(\underline{p},\underline{q}) \in (U - DE, U)$, inoltre il parametro DE è "arbitrario", come discusso prima della (3.8).

Tuttavia il problema dell' ortodicità o meno dell' insieme statistico $\mathcal{E}$ i cui elementi sono parametrizzati da U e V via la (5.7) è "solo" la questione se sia o no valida la (5.6) e tale problema non è di per se logicamente o matematicamente legato ad alcuna proprietà della dinamica microscopica

La relazione fra la ortodicità di un insieme statistico e le ipotesi sulla dinamica microscopica che la garantirebbero a priori (e.g. ipotesi ergodica etc.) sarà ripresa più in dettaglio a conclusione del §6.

Se esistessero insiemi statistici "ortodici" allora ognuno di essi fornirebbe un modello microscopico di termodinamica classica; ovviamente, se esistessero più insiemi statistici ortodici possibili, dovrebbe anche accadere che i vari modelli microscopici di termodinamica che essi forniscono siano equivalenti, ossia forniscano la stessa espressione della S in termine delle altre grandezze termodinamiche che risulterebbero così definibili in termini meccanici, in modo non ambiguo. Questa è una delle questioni affrontate nella teoria degli Insiemi Statistici.

Si intravede qui che se si tenta di abbandonare il punto di vista fondamentale in cui si cerca di fondare la termodinamica sulla meccanica, si dovranno ugualmente affrontare enormi problemi quali quello della non ambiguità della termodinamica da associare ad un dato sistema. Questo problema è risolto in vari casi importanti ma si è assai lontani dall' essere sicuri che tali casi (insiemi statistici microcanonico, canonico, gran canonico etc., si veda Insiemi Statistici) esauriscano tutti i possibili. In ultima analisi una completa soluzione di questo problema potrebbe rivelarsi equivalente

al problema della fondazione dinamica della termodinamica che ci si illude, invece, di accerchiare.

§6 Modelli termodinamici. Gli insiemi statistici microcanonico e canonico e loro relazione con l' ipotesi ergodica

Il problema della esistenza di insiemi statistici (cioè di insiemi di distribuzioni stazionarie definite sullo spazio delle fasi) che forniscano modelli microscopici di termodinamica almeno nel limite in cui il volume V occupato dal sistema diviene infinito, ma la densità media e l' energia cinetica media restano fisse, "limite termodinamico", fu risolto da Boltzmann nello stesso lavoro citato sopra.

Ivi Boltzmann dimostrò che i due seguenti insiemi statistici, descritti nelle righe che seguono e detti insieme "microcanonico" e "canonico", sono "ortodici" ossia definiscono un modello microscopico di termodinamica in cui l' energia cinetica media per particella è proporzionale alla temperatura assoluta (c.f.r. seguito e §5).

1) *L' "insieme microcanonico"*

così chiamato da Gibbs mentre Boltzmann lo chiamò con il nome, ancora famoso ma subito caduto in disuso, di "ergodo"): è costituito dalle misure stazionarie μ parametrizzate da due parametri U = energia totale e V = volume a disposizione del sistema in modo che (si veda (5.2)):

$$\begin{aligned} &\mu(\Delta) = 1/\mathcal{N}(U,V) \qquad &&\text{se } U - DE \le E(\Delta) \le U \\ &\mu(\Delta) = 0 &&\text{altrimenti} \end{aligned} \tag{6.1}$$

con:

$$\mathcal{N}(U,V) = \sum_{U-DE \le E(\Delta) \le U} 1 = \{\text{numero delle cellette } \Delta \text{ di energia } E(\Delta) \in (U-DE, U)\} \tag{6.2}$$

ove la quantità DE deve essere una quantità, eventualmente dipendente da V, "macroscopicamente trascurabile" affinché si possa considerare che tutte le cellette con energia fra $U - DE$ ed U abbiano la stessa energia da un punto di vista delle misurazioni macroscopiche.

L' importanza dell' insieme microcanonico nella relazione fra la termodinamica classica e l' ipotesi atomica è illustrata dall' argomento che conduce alla (3.8) e che propone l' insieme microcanonico come candidato naturale per un esempio di insieme statistico ortodico: tuttavia come discusso al §5 l' argomento che conduce alla (3.8) e ne motiva l' introduzione non è in alcun senso una dimostrazione della ortodicità dell' insieme microcanonico

(la quale, convien ripetere, da un punto di vista matematico non dipende da ipotesi sulla dinamica microscopica).

In base alla definizione generale del §5 di insieme statistico ortodico, cioè originante un modello di termodinamica, si possono definire la "temperatura assoluta" e la "entropia" di ogni elemento μ ("stato macroscopico") dell' insieme, in modo che la temperatura T sia proporzionale all' energia cinetica media. Boltzmann mostrò che tali funzioni T ed S sono date dalle celebri relazioni:

$$T = \frac{2}{3k}\frac{T(\mu)}{N}\ , \qquad S(\mu) = k \log \mathcal{N}(U, V) \tag{6.4}$$

ove k, "costante di Boltzmann", è una costante universale da determinare empiricamente per confronto fra teoria ed esperienza (come si vedrà in seguito si trova: $k = 1.38\ 10^{-16}$ erg/°K); il fattore $2/3$ è convenzionale e questa scelta semplifica alcune formule successive, oltre alla seconda delle (6.4).

L' affermazione che (6.1) e (6.2) forniscono un "modello microscopico di termodinamica" nel limite termodinamico $V \to \infty$, $U \to \infty$, $N \to \infty$ in modo che $u = U/N$, $v = V/N$ restino costanti, è da interpretarsi nel senso seguente.

A partire da (6.1)/(6.4) si calcolano, si veda (5.5):

$$\begin{aligned} &u = U/N = \text{ "energia specifica" } v = V/N = \text{ "volume specifico"} \\ &T = 2T(\mu)/3kN = \text{ "temperatura" } s = S(\mu)/N = \text{ "entropia"} \\ &P = P(\mu) = \text{ "pressione".} \end{aligned} \tag{6.5}$$

Poiché u, v determinano $\mu \in \mathcal{E}$ si potranno evidentemente esprimere T, P, s in funzione di u, v a mezzo di funzioni $T(u, v)$, $P(u, v)$, $s(u, v)$ che ammetteremo che tendano a funzioni limite, nel limite per $V \to \infty$, con u, v fissi.

Dire che (6.1) e (6.2) forniscono un modello di termodinamica significa (si veda anche §5) che tali funzioni verificano le stesse relazioni intercorrenti fra le omonime grandezze nella termodinamica classica e cioè:

$$\mathrm{d}u = T\,\mathrm{d}s + P\,\mathrm{d}v \tag{6.6}$$

La (6.6) va letta così: se si fa variare di poco lo stato μ definito da (6.1), le corrispondenti variazioni di u, s, v verificano le (6.6), cioè il "secondo principio della termodinamica". per una discussione e dimostrazione di (6.4), (6.6) si veda la voce "Insiemi Statistici".

2) *L' "insieme canonico"*

così chiamato da Gibbs, mentre Boltzmann lo chiamò "olodo", è costituito dall' insieme delle distribuzioni stazionarie parametrizzate da due parametri, β e $v = V/N$, via la definizione:

$$\mu(\Delta) = (\exp -\beta E(\Delta))/Z(\beta, V) \tag{6.7}$$

con

$$Z(\beta, V) = \sum_{\Delta} \exp -\beta E(\Delta) \tag{6.8}$$

Boltzmann dimostrò l' ortodicità di questo insieme statistico facendo vedere che temperatura ed entropia possono essere definite da:

$$T = 2T(\mu)/3kN = 1/k\beta \qquad S = -k(\beta U - \log Z(\beta, V)) \tag{6.9}$$

ove k è una costante universale da determinare empiricamente.

L' affermazione che (6.7), (6.9) forniscono un modello microscopico di termodinamica classica nel limite termodinamico $V \to \infty$, $V/N \to v$, $\beta =$ costante ha lo stesso significato descritto nel caso precedente.

Per la discussione e dimostrazione della ortodicità dell' insieme canonico si veda la vode "Insiemi Statistici".

Le relazioni (6.6) valgono, come si è detto, nei due casi di insiemi statistici considerati sopra e, quindi, ciascuno di questi insiemi fornisce un modello microscopico di termodinamica classica.

Poiché l' entropia, la pressione, la temperatura etc. sono in entrambi i casi esplicitamente esprimibili in termini di due parametri indipendenti (u e v oppure β e v) si sarà in grado di calcolare l' " equazione di stato" (cioè il legame fra P, v e T) in termini delle proprietà microscopiche del sistema, almeno in linea di principio: è questo un enorme progresso rispetto alla termodinamica classica, ove l' equazione di stato ha sempre carattere fenomenologico, ossia è una relazione che può essere dedotta unicamente dall' esperienza.

È chiaro, però, che i modelli di termodinamica sopra descritti, per essere accettabili, dovranno rispondere al fondamentale requisito di definire non solo una termodinamica possibile (ossia non in contrasto con i principi della termodinamica espressi dalle (6.6)), ma anche di definire "la" termodinamica del dato sistema, quella cioè accessibile sperimentalmente.

Per questo è, anzitutto, necessario che i due modelli di termodinamica coincidano, ossia conducano alle stesse relazioni fra le grandezze termodinamiche fondamentali u, v, T, P, s, ma è anche necessario che tali modelli siano in accordo con le osservazioni sperimentali.

A priori non ci sono, però, ragioni perché queste due proprietà siano valide.

Ci si può qui riallacciare, per approfondire le questioni legate alla (3.8) e produrre una giustificazione della validità dell' insieme microcanonico

come modello di termodinamica, all' ipotesi ergodica che a volte è invocata a questo punto al fine di garantire a priori il successo delle verifiche di consistenza teorica e sperimentale la cui necessità logica si è ora messa in evidenza.

Nei paragrafi precedenti si è visto infatti, c.f.r. §3, come la distribuzione microcanonica potesse essere giustificata come descrivente stati di equilibrio macroscopico sulla base della ipotesi ergodica e di una certa proprietà di continuità: in quella analisi, che conduceva alla (3.8), non si è però tenuto conto alcuno delle scale di tempo coinvolte nella discussione. La loro importanza è stata messa in luce nel §4: se la (3.8) fosse vera ma il valore medio temporale di f raggiungesse il suo valore limite, dato dal membro di destra della (3.8), in un tempo "superastronomico" quale quello dato dalla (4.3) allora la (3.8) avrebbe, ovviamente, scarso valore pratico.

Riassumendo, dunque, per la deduzione della (3.8) e quindi per la giustificazione a priori della connessione fra l' insieme microcanonico e l' insieme degli stati di equilibrio termodinamico macroscopico si incontrano tre difficoltà.

La prima è la verifica dell' ipotesi ergodica.

La seconda è che, anche accettando l' ipotesi ergodica per la distribuzione stazionaria sulle cellette di energia fissata microscopicamente (cioè a meno di δE), si deve superare la difficoltà della non ergodicità degli elementi μ dell' insieme microcanonico che fa si che la relazione (5.4) non sia necessariamente vera (si ricordi che la non ergodicità è dovuta al fatto che nell' insieme microcanonico l' energia fluttua di una quantità piccola ma macroscopica $DE \gg \delta E$, e quindi $\mu(\Delta) = 0$ per cellette Δ con energie diverse e quindi appartenenti a cicli diversi della dinamica S).

La terza è che, comunque, sembrerebbe che sia necessario attendere tempi enormi (enormemente più lunghi della età dell' universo, nei casi più interessanti), prima che le fluttuazioni dei valori medi delle osservabili si stabilizzino sul valore limite di equilibrio.

Le tre difficoltà si risolverebbero se si supponesse simultaneamente:
i) le cellette di data energia (fissata microscopicamente) fanno parte di un solo ciclo di evoluzione: "ipotesi ergodica", c.f.r. §3;
ii) le osservabili macroscopiche interessanti sono praticamente costanti sulle cellette di una data componente ergodica del sistema (corrispondente ad un valore macroscopico E dell' energia);
iii) il comune valore medio che le osservabili macroscopiche interessanti hanno sulle cellette di energia E varia poco al variare della energia E fra $U - DE$ ed U se U e DE sono due valori macroscopici con $U \gg DE$ (pur essendo $DE \gg \delta E$), eccezion fatta per una piccola frazione di cellette, trascurabile nel limite termodinamico.

Le ipotesi i) e iii), c.f.r. §3, fanno vedere che i valori medi delle osserv-

abili macroscopiche possono essere calcolati usando indifferentemente una componente ergodica di una data distribuzione microcanonica μ o la intera distribuzione microcanonica μ data.

La ipotesi ii) permette di dire che il tempo necessario perché il valore medio di un' osservabile interessante sia raggiunto, se calcolato sull' evoluzione di un particolare stato microscopico Δ, è di gran lunga inferiore al tempo di ricorrenza (troppo lungo per essere interessante).

Accettare le ipotesi i), ii), iii) implica (per il significato fisico che u, p e v acquistano) che l' insieme microcanonico *deve* fornire un modello di termodinamica nel senso che $\mathrm{d}u+p\,\mathrm{d}v$ deve ammettere un fattore integrante (da identificare con la temperatura assoluta). Il fatto poi che tale fattore risulti proporzionale alla energia cinetica media è, da questo punto di vista (e *nel solo caso della meccanica Statistica classica*) una conseguenza, si veda in proposito la voce Insiemi Statistici.

Si può qui notare che le ipotesi ii) e iii) sono ipotesi che, almeno ad un livello qualitativo, non coinvolgono esplicitamente le proprietà della evoluzione dinamica del sistema: sono, come si suol dire, "proprietà di equilibrio" del sistema.

Ed è molto ragionevole pensare che siano verificate per la maggior parte dei sistemi che si incontrano nelle applicazioni, perché in molti casi sono effettivamente verificabili, e talvolta con totale rigore matematico.

Dunque la ipotesi più profonda è la i): ed è in questo senso che a volte, impropriamente, si dice che l' ipotesi ergodica è il fondamento teorico dell' uso dell' insieme microcanonico quale modello microscopico dell' insieme degli stati di equilibrio di un sistema.

La improprietà della locuzione sta nel fatto che la i) può essere di gran lunga indebolita senza modificare la conclusione sull' interpretazione dell' insieme microcanonico. Ad esempio si potrebbe richiedere che la media temporale delle *sole* osservabili macroscopicamente interessanti abbia lo stesso valore su qualsiasi ciclo (o sulla maggior parte dei cicli) di data energia, pur ammettendo che sulla superficie di data energia esistano vari cicli diversi (sui quali solo osservabili non macroscopicamente interessanti assumono valori medi diversi).

Inoltre la locuzione menzionata è impropria perché, se anche la si accetta, non si può prescindere dalla verifica delle ii) e iii) e, in particolare, di tali ipotesi è necessaria una verifica quantitativa non potendo essere evidentemente soddisfacente limitarsi alla loro sola verifica qualitativa. Si potrebbe invero dubitare che il tempo di "raggiungimento dell' equilibrio" possa realmente scendere dai tempi di ricorrenza (superastronomici) ai tempi osservati sperimentalmente (usualmente di pochi microsecondi).

Per quello che riguarda poi l' insieme canonico, il suo uso potrebbe essere giustificato semplicemente mostrando che produce gli stessi risultati che si

ottengono attraverso l' uso dell' insieme microcanonico, almeno nel limite termodinamico.

Ma, come si è detto, l' ipotesi ergodica (rafforzata o no dalle ipotesi ii) e iii) sopra descritte) è tecnicamente troppo difficile da studiare e per questo si è tentato di affrontare il problema della costruzione di modelli microscopici per la termodinamica evitando di risolvere, seppure parzialmente, il problema ergodico.

La proposta è semplicemente di dimostrare che tutti gli insiemi statistici ortodici (e "ragionevoli") producono la *stessa* termodinamica macroscopica (ad esempio la stessa equazione di stato per un dato sistema) e di considerare questa proprietà, notevolissima e sorprendente di per se, come sufficiente per postulare, via il "principio di ragion sufficiente", che le equazioni di stato di un sistema sono calcolabili dalle proprietà microscopiche (ossia dalla hamiltoniana) valutando i valori medi delle osservabili fondamentali tramite le distribuzioni degli insiemi microcanonico o canonico o più in generale di un qualsiasi insieme statistico ortodico: è questo il punto di vista di solito attribuito a Gibbs e sul quale si basano quasi tutte le trattazioni della termodinamica statistica.

Ben si comprende come un tale punto di vista fosse insoddisfacente per Boltzmann, che ambiva a ridurre la termodinamica alla meccanica senza l' introduzione di alcun nuovo postulato: d'altra parte il punto di vista pragmatico di Gibbs è comprensibile alla luce della necessità di trarre tutte le conseguenze applicative dalla meravigliosa scoperta della possibilità di calcolare univocamente le grandezze termodinamiche in termini delle proprietà meccaniche del modello atomico della materia.

Da qualche decennio, ad un secolo dalla nascita di queste teorie, si sente di nuovo la necessità unitaria di dedurre la termodinamica dalla meccanica senza l' artificioso postulato a priori che la termodinamica sia descritta dagli elementi degli insiemi statistici ortodici (postulato reso possibile dalla indipendenza, già menzionata e che vedremo (cfr. "Insiemi Statistici"), dei risultati dal particolare insieme statistico che si considera).

Il problema ergodico e la dinamica statistica sono così tornati ad essere al centro della ricerca, stimolando alcuni nuovi risultati interessanti.

Boltzmanntentò di affrontare il problema della giustificazione degli insiemi statistici canonico e microcanonico seguendo anche una strada diversa, oltre a quella dello studio del problema ergodico e delle ipotesi i), ii), iii) ora descritte, giungendo alla deduzione della "equazione di Boltzmann", c.f.r. anche il §7 successivo per una analisi a grandi linee di questa equazione, che si è poi rivelata preziosa addirittura per le applicazioni più tecniche pur presentando vari aspetti concettualmente insoddisfacenti.

§7 L' avvicinamento all' equilibrio e l' equazione di Boltzmann. Ergodicità

ed irreversibilità

Come si è detto gli stati di equilibrio possono essere identificati con gli elementi degli insiemi statistici ortodici (microcanonico, canonico, gran canonico, etc...).
Tuttavia non è chiaro attraverso quale meccanismo un sistema meccanico inizialmente in una situazione di non equilibrio raggiunga l' equilibrio.
Si è visto che l' ipotesi ergodica di per sé non è sufficiente a spiegare come mai un sistema raggiunga l' equilibrio in tempi, di solito, relativamente brevi.

Boltzmann ha fornito di questo meccanismo un modello, contestatissimo fin dalla sua formulazione come le altre sue intuizioni, che alcuni, credo a torto, vedono come il suo maggior contributo alla scienza.
Questo modello ha valdità limitata a sistemi di densità così bassa da essere pensabili come gas rarefatti e mostra come in concreto possa avvenire che le ipotesi i), ii), iii) del §6 siano, agli effetti pratici, soddisfatte in tali sistemi e come sia possibile che le osservabili interessanti raggiungano i loro valori medi su scale di tempo osservabili anziché sulle assurdamente lunghe scale dei tempi di ricorrenza.
Si immagina dunque che il sistema consista di N particelle identiche (per semplicità), ciascuna delle quali è descritta dall' impulso $\underline{p}$ e dalla posizione $\underline{q}$.
Queste particelle si muovono come se fossero libere eccetto che, di tanto in tanto, subiscono collisioni.
Supponendo che tali particelle siano sfere rigide con raggio R (di nuovo solo per semplicità) e dotate di velocità media $\bar{v}$, l' ipotesi di bassa densità è che la densità $\rho = N/V$ sia tale che:

$$\rho R^3 \ll 1 \tag{7.1}$$

che significa che è molto improbabile che ci siano due particelle a distanza dell' ordine di R, cioè "in collisione".
Al tempo stesso si richiede che il numero di collisioni che ogni particella subisce per unità di tempo sia non nullo. Evidentemente tale numero è dell' ordine di:

$$\rho R^2 \overline{v} \tag{7.2}$$

Quindi la situazione limite in cui il gas è molto rarefatto ma, ciononostante, il numero di urti di ogni particella, per unità di tempo, è non trascurabile è descritta da:

$$\begin{aligned} &R \to 0, \rho \to \infty \quad \text{in modo che} \\ &\rho R^3 \to 0, \rho R^2 \overline{v} = w = \text{ quantità fissa} \end{aligned} \tag{7.3}$$

La quantità $\tau = 1/w$ è il "tempo di volo" medio fra due collisioni mentre il "libero cammino medio" è $\tau\overline{v} = 1/\rho R^2$.

La situazione limite, che si raggiunge facendo tendere R a 0 e ρ a ∞ secondo la (7.3), è chiamata il "limite di Grad–Boltzmann": nella situazione considerata da Boltzmann si suppone di essere "vicini" a tale limite, cioè si suppone che $\rho R^3 \ll 1$ e $\rho R^2\overline{v} = w > 0$.

Sia allora $f(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}$ il numero di particelle che si trovano nella celletta $Q = \mathrm{d}\underline{p}\,\mathrm{d}\underline{q}$ dello spazio delle fasi che descrive gli stati di una *singola* particella.

Boltzmann osserva che f può variare nel tempo sia in virtù delle collisioni sia del fatto che le particelle si muovono nello spazio.

Se ε è un intervallo di tempo prefissato, le particelle che si trovano ad un certo istante t nella celletta Q sono in numero di:

$$\begin{aligned} f(\underline{p},\underline{q},\underline{t})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q} &= f(\underline{p},\underline{q}-\varepsilon\underline{p}/m,t-\varepsilon)\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q} + \\ &+ \sum_{Q',Q''} \text{(numero di particelle in } Q' \text{ che per unità di tempo collidono} \\ &\qquad \text{con particelle in } Q'' \text{ producendo particelle in } Q_1, Q_2 \text{ con } Q_1 \equiv Q) - \\ &- \sum_{Q',Q''} \text{(numero di particelle in } Q_1 \equiv Q \text{ che per unità di tempo collidono} \\ &\qquad \text{con particelle in } Q_2 \text{ producendo particelle in } Q', Q'') \end{aligned} \tag{7.4}$$

Se ora si considera la collisione che a due particelle in Q', Q'' ne associa altre due in Q_1, Q_2 si dovrà avere (per la conservazione del momento e dell' energia nell' urto):

$$\underline{p}' + \underline{p}'' = \underline{p}_1 + \underline{p}_2 \quad , \quad \underline{p}'^2 + \underline{p}''^2 = \underline{p}_1^2 + \underline{p}_2^2 \tag{7.5}$$

e il numero di collisioni che portano da p', p'' a $\underline{p}_1$, $\underline{p}_2$ può essere espresso in termini della nozione di "sezione d'urto" di collisione.

Introducendo:

$$\begin{aligned} f(\underline{p}',\underline{q})\,\mathrm{d}\underline{p}'\,\mathrm{d}q &= \text{numero di particelle di impulso } \underline{p}', \\ &\quad \text{a meno di } \mathrm{d}\underline{p}' \text{ nel cubetto } \mathrm{d}\underline{q} = \\ &\quad \text{"numero di centri di collisione"} \\ f(\underline{p}'',\underline{q})\,\mathrm{d}\underline{p}'' &= \text{densità delle particelle di impulso } \underline{p}'', \\ &\quad \text{a meno di } \mathrm{d}\underline{p}'', \text{ nel punto } \underline{q} = \\ &\quad = \text{"densità delle particelle che} \\ &\quad \text{possono subire un urto"} \\ \sigma(\underline{p}',\underline{p}'';\underline{p},\underline{p}_2) &= \text{sezione d'urto per la} \\ &\quad \text{collisione in questione} \end{aligned}$$

si ha che il "volume di collisione" associato ad un solo centro di collisione è (ricordando che la velocità relativa nella collisione è $|\underline{p}' - \underline{p}''|/m$, $\underline{p} \equiv \underline{p}_1$):

$$(|\underline{p}' - \underline{p}''|/m)\sigma(\underline{p}',\underline{p}'',\underline{p},\underline{p}_2)$$

che per definizione di sezione d'urto dà, se moltiplicato per la densità delle particelle di impulso $\underline{p}''$, il numero per unità di tempo delle collisioni che le particelle di impulso $\underline{p}''$ subirebbero con particelle di impulso $\underline{p}'$ se ci fosse una sola particella di impulso $\underline{p}'$ in $\underline{q}$.

Dunque il numero totale di collisioni da Q', Q'' a Q_1, Q_2 è, per unità di tempo:

$$f(\underline{p}'', \underline{q})\, \mathrm{d}\underline{p}'' \frac{|\underline{p}' - \underline{p}''|}{m} \sigma(\underline{p}', \underline{p}'', \underline{p}, \underline{p}_2) f(\underline{p}', \underline{q})\, \mathrm{d}\underline{p}'\, \mathrm{d}\underline{q} \tag{7.6}$$

ovviamente simmetrica in $\underline{p}'$, $\underline{p}''$ sebbene dedotta trattando $\underline{p}'$ e $\underline{p}''$ in modo asimmetrico.

Per un analogo argomento il numero di collisioni "inverse" è:

$$f(\underline{p}, \underline{q}) f(\underline{p}_1, \underline{q})\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{p}_2\, \mathrm{d}\underline{q} \frac{|\underline{p} - \underline{p}_2|}{m} \sigma(\underline{p}_1, \underline{p}_2; \underline{p}', \underline{p}'') \tag{7.7}$$

Si osserva poi che dalle (7.5) segue che:

$$\begin{aligned} &\mathrm{d}\underline{p}'\, \mathrm{d}\underline{p}'' = \mathrm{d}\underline{p}\, \mathrm{d}\underline{p}_2 \qquad (\text{"teorema di Liouville" }) \\ &\underline{p}' - \underline{p}'' = \underline{p} - \underline{p}_2 \end{aligned} \tag{7.8}$$

e inoltre la sezione d'urto, come in generale per collisioni governate da forze centrali, dipende esclusivamente dall' "angolo di deflessione" fra $(\underline{p}' - \underline{p}'')$ e $(\underline{p} - \underline{p}_2)$ e dalla velocità relativa $|\underline{p}' - \underline{p}''|/m$, ed è proporzionale all' angolo solido $\mathrm{d}\Omega$ in cui si immagina diretto $(\underline{p} - \underline{p}_2)$ rispetto a $(p' - p'')$.

Si noti a questo proposito che i dati finali della collisione, cioè $\underline{p}$ e $\underline{p}_2$, non determinano $\underline{p}'$ e $\underline{p}''$ via le (7.5) ma lasciano arbitraria la direzione $\mathrm{d}\Omega$ di $\underline{p}' - \underline{p}'')$.

Si porrà allora $\sigma(\underline{p}', \underline{p}'', \underline{p}, \underline{p}_2) = \sigma(\underline{\omega}, |\underline{p}' - \underline{p}''|)\, \mathrm{d}\underline{\omega} = \sigma(\underline{\omega})\, \mathrm{d}\underline{\omega}$ ove l' ultima relazione è valida solo nel caso in cui l' interazione delle sfere è supposta una interazione fra sfere rigide (e dalla teoria delle collisioni risulta anzi che $\sigma(\underline{\omega})$ è indipendente da $\underline{\omega}$: $\sigma(\underline{\omega}) = 4\pi R^2$).

Dunque (7.6), (7.7), (7.8) permettono di riscrivere le (7.6) e (7.7) come:

$$\begin{aligned} &f(\underline{p}', \underline{q}) f(\underline{p}'', \underline{q})\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{p}_2\, \mathrm{d}\underline{q}\, \mathrm{d}\underline{\omega} (|\underline{p}' - \underline{p}''|/m) \sigma(\underline{\omega}) \\ &f(\underline{p}, \underline{q}) f(\underline{p}_2, \underline{q})\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{p}_2\, \mathrm{d}\underline{q}\, \mathrm{d}\underline{\omega} (|\underline{p}' - \underline{p}''|/m) \sigma(\underline{\omega}) \end{aligned} \tag{7.9}$$

dove, dati $\underline{p}$, $\underline{p}_2$, i vettori $\underline{p}'$, $\underline{p}''$ si calcolano dalle (7.5) e dalla informazione che l' angolo solido fra $\underline{p} - \underline{p}_2$ e $\underline{p}' - \underline{p}''$ è $\underline{\omega}$.

Introducendo la (7.9) nella (7.3) e dividendo per ε si trova l' "equazione di Boltzmann":

$$\begin{aligned} &\frac{\partial f}{\partial t}(\underline{p}, \underline{q}) + \frac{\underline{p}}{m} \cdot \frac{\partial f}{\partial \underline{q}}(p, q) = \int (|\underline{p} - \underline{q}_2|/m) \sigma(\underline{\omega})\, \mathrm{d}\underline{\omega}\, \mathrm{d}\underline{p}_2 \\ &(f(\underline{p}', \underline{q}) f(\underline{p}'', \underline{q}) - f(\underline{p}, \underline{q}) f(\underline{p}_2, \underline{q})) \end{aligned} \tag{7.10}$$

Nella (7.10) si suppone che $\underline{q}$ vari in tutto lo spazio: però i casi più interessanti si hanno quando il sistema è confinato in un dato volume V.

In tal caso la (7.10) va corredata di opportune condizioni al contorno che dipendono da come le particelle urtano contro le pareti. Poiché la discussione delle condizioni al contorno è delicata la eviteremo e nei casi di sistemi confinati supporremo per semplicità che al bordo valgano "condizioni periodiche". Questo significa che si immagina il volume V come un cubo le cui facce opposte sono identificate, il che significa che ogni particella che collide con una faccia del cubo emerge, dopo la collisione, dalla faccia opposta e con la stessa velocità. Per una trattazione più esauriente del problema delle condizioni al contorno si veda la voce "Teorie Cinetiche".

È chiaro che la (7.10) è una approssimazione perché in essa si trascurano:
i) la possibilità di collisioni multiple;
ii) la possibilità che le particelle che si trovano in una data celletta non si comportino indipendentemente le une dalle altre (come implicitamente assunto nella derivazione della (7.10)) e anzi fra di loro si costruiscano, al crescere del tempo, correlazioni fra posizioni e velocità che rendono più probabili certi urti rispetto ad altri, ovvero collisioni multiple rispetto a quelle binarie.

Questi effetti dovrebbero sparire nel limite di Grad-Boltzmann (7.3), purché siano assenti all' istante iniziale: tale congettura è chiamata "congettura di Grad" sulla validità della "stosszahlansatz" (parola quest'ultima che, per motivi tradizionali, denota appunto la mancanza di correlazioni fra i moti delle varie particelle in vari istanti).

Ritornando all' equazione di Boltzmann (e posponendo l' analisi delle ipotesi fondamentali i) e ii) sopra discusse), la irreversibilità da essa implicata può essere dimostrata sulla base delle seguenti osservazioni.

Moltiplicando membro a membro la (7.10) per 1, $\underline{p}$, $\underline{p}^2/2m$ ovvero per $(1+\log f(\underline{p},\underline{q}))$ e integrando su $\underline{p}$ e $\underline{q}$ si trova che le quantità:

$$N=\int f(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}, \qquad \underline{P}=\int \underline{p} f(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}$$
$$T=\int (\underline{p}^2/2m) f(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}\,, \qquad H=-\int f(\underline{p},\underline{q})\log f(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q} \tag{7.11}$$

verificano, nell' ipotesi che $f(\underline{p},\underline{q})\to 0$ velocemente per $\underline{p}\to\infty$, e anche $f(\underline{p},\underline{q})\to 0$ per $\underline{q}\to\infty$ velocemente, (ovvero, se $\underline{q}$ varia in un contenitore prefissato e se f verifica opportune condizioni al contorno sulle $\underline{q}$), le relazioni:

$$\frac{\mathrm{d}N}{\mathrm{d}t}=\frac{\mathrm{d}\underline{P}}{\mathrm{d}t}=\frac{\mathrm{d}T}{\mathrm{d}t}=0$$
$$\frac{\mathrm{d}H}{\mathrm{d}t}=\frac{1}{4}\int \frac{|\underline{p}-\underline{p}_2|}{m}\sigma(\underline{\omega})\,\mathrm{d}\underline{\omega}(f(\underline{p}',\underline{q})f(\underline{p}'',\underline{q})-f(\underline{p},\underline{q})f(\underline{p}_2,\underline{q}))\cdot$$

$$\cdot (\log f(\underline{p}', \underline{q}) f(\underline{p}'', \underline{q}) - \log f(\underline{p}, \underline{q}) f(\underline{p}_2, \underline{q})) \, \mathrm{d}\underline{p} \, \mathrm{d}\underline{p}_2 \, \mathrm{d}\underline{q} \geq 0 \quad (7.12)$$

come mostra un semplice calcolo nel quale gioca un ruolo essenziale la simmetria fra $\underline{p}, \underline{p}_2$ e $\underline{p}', \underline{p}''$ del secondo membro della (7.10) e inoltre si usa la $\mathrm{d}\underline{p} \, \mathrm{d}\underline{p}_2 = \mathrm{d}\underline{p}' \, \mathrm{d}\underline{p}''$ e le proprietà $\log x + \log y = \log xy$ e $(x-y)(\log x - \log y) \geq 0$.

Quindi mentre dalle prime tre relazioni delle (7.12) seguono cinque leggi di conservazione (del numero di particelle, dell' impulso e dell' energia (cinetica)), dalla quarta, detta "teorema H di Boltzmann", segue manifestamente la irreversibilità della evoluzione.

Anzi la (7.12) mostra che possono essere stati di equilibrio solo quelli per cui:

$$f(\underline{p}', \underline{q}) f(\underline{p}'', \underline{q}) = f(\underline{p}, \underline{q}) f(\underline{p}_2, \underline{q}) \quad (7.13)$$

se $\underline{p}$, $\underline{p}_2$, $\underline{p}'$, $\underline{p}''$ verificano le (7.4).

Dalla (7.13) e dalla arbitrarietà di $\underline{p}$, $\underline{p}_2$, $\underline{p}'$, $\underline{p}''$ si deduce con un semplice argomento che tralasciamo che:

$$f(\underline{p}, \underline{q}) = \frac{\rho(\underline{q}) e^{-\beta(\underline{q})(\underline{p} - \underline{p}_0(q))^2/2m}}{(2\pi\rho(\underline{q})m)^{3/2}} \quad (7.14)$$

ove $\beta(\underline{q}), \underline{p}_0(\underline{q})$ e $\rho(\underline{q})$ sono funzioni arbitrarie ed il fattore ultimo sotto il sego di radice è stato introdotto per comodità, in modo che $\rho(\underline{q})$ possa essere interpretato come densità nel punto q: $\rho(\underline{q}) = \int f(\underline{p}, \underline{q}) \, \mathrm{d}\underline{p}$.

Dunque gli stati di equilibrio hanno necessariamente la forma (7.14).

Considerando poi il caso semplice di un sistema racchiuso in un contenitore cubico con condizioni periodiche al contorno è facile vedere che, se f verifica (7.10), (7.14) e $\partial f/\partial t = 0$ (cioè è stazionaria) allora deve essere necessariamente vero che $\beta(\underline{q})$, $\rho(\underline{q})$ e $\underline{p}_0(\underline{q})$ sono indipendenti da $\underline{q}$.

Infatti se f ha la forma (7.14) il membro di destra di (7.10) è nullo e, quindi, la $\partial f/\partial t = 0$ implica $\underline{p} \cdot \partial f/\partial \underline{q} = 0$ e cioè, denotando $\hat{f}(\underline{p}, \underline{k})$ la trasformata di Fourier di f rispetto a q, implica $\underline{p} \cdot \underline{k} \hat{f}(\underline{p}, \underline{k}) = 0$: dunque se $\hat{f}(\underline{p}, \underline{k})$ è continua in $\underline{p}$, deve essere $\hat{f}(\underline{p}, \underline{k}) = 0$ per $\underline{k} \neq 0$ e questo significa che f è indipendente da $\underline{q}$ e quindi $\beta(\underline{q})$, $\rho(\underline{q})$, $\underline{p}_0(\underline{q})$ sono costanti.

Dunque il teorema H non solo mostra che il sistema evolve irreversibilmente, ma anche che evolve verso la distribuzione di "Maxwell-Boltzmann" che altro non è che un elemento μ dell' insieme canonico in un sistema in cui l' interazione fra le particelle sia così piccola che l' energia totale del sistema possa essere identificata con l' energia cinetica, che è appunto quel che accade nel limite di Boltzmann-Grad e quel che ci si deve attendere che avvenga nei gas rarefatti: e i parametri β, ρ, $\underline{p}_0$ di questa distribuzione sono univocamente determinati dai dati iniziali via le leggi di conservazione (7.12).

Ed è spontaneo pensare che il teorema H sia, per i gas rarefatti, la versione microscopica del secondo Principio della termodinamica che prevede l' accrescimento dell' entropia (da identificarsi come proporzionale a H).

Tuttavia è bene sottolineare che il teorema H è in palese contrasto con le proprietà di reversibilità delle equazioni di Hamilton e, dunque, come già osservato non ne può essere una "conseguenza" nel senso ingenuo della parola.

Proprio per questo è essenziale capire se questo contrasto fra irreversibilità macroscopica e reversibilità microscopica possa essere superato.

Sulla presunta incompatibilità fra queste due proprietà Boltzmann subì violente critiche e creò l' ipotesi e (quindi) la teoria ergodica nel tentativo di rispondere ad esse su una base teorica solidamente basata sui fondamenti della meccanica, anziché sulla abbastanza oscura stosszahlansatz.

Per affrontare la questione si possono assumere due punti di vista, "sovrapposti" l' uno all' altro da Boltzmann nel suo brillante e mal compreso tentativo di difesa della sua teoria e del suo teorema H.

Il primo è che valga l' ipotesi ergodica (nel senso i) del §6 rafforzato da ii) e iii)) e quindi $\frac{\mathrm{d}H}{\mathrm{d}t} \geq 0$ possa essere solo approssimativamente vera nel senso che sarebbe vera per la maggior parte del tempo: quando la celletta $S^k\Delta$ che rappresenta lo stato microscopico all' istante $k\tau$ percorre la maggior parte del ciclo ergodico di data energia, (cioè la parte nella quale le osservabili macroscopiche sono pure praticamente costanti). La relazione $\mathrm{d}H/\,\mathrm{d}t \geq 0$ diverrebbe poi falsa quando $S^k\Delta$ esce da tale regione.

Quest'ultima circostanza però può avvenire, per i sistemi realmente macroscopici, con frequenza temporale ampiamente più lunga delle più lunghe scale di tempo astronomiche (si veda il §4, (4.3)).

Dunque il sistema evolverebbe in modo praticamente irreversibile (e simmetricamente nel tempo!) e la reversibilità si manifesterebbe macroscopicamnte su scale di tempo "al di là dell' eternità", cioè di vari ordini di grandezza superiori all' età dell' Universo, già per sistemi quali un gas a condizioni normali contenuto in un recipiente delle dimensioni di una stanza (o di una scatoletta).

Un sistema posto inizialmente in una condizione "atipica", ad esempio occupante uniformemente la metà del contenitore, si espanderebbe ad occupare l' intero contenitore e poi continuerebbe ad evolvere senza "mai più" ritornare ad occupare la metà iniziale del contenitore. Ovviamente se un "demonietto", intervenendo dopo poco tempo dall' istante iniziale, invertisse le velocità di tutte le particelle del sistema, il sistema procederebbe a ritroso nel tempo ritornando in breve (e per un tempo brevissimo) ad occupare solo metà del recipiente e poi evolverebbe ancora occupando tutto il contenitore evolvendo all' equilibrio esattamente come avrebbe fatto se le velocità non fossero state invertite (e per giunta secondo una legge di

evoluzione descritta approssimativamente dall' equazione di Boltzmann). Questa inversione del moto con produzione di una situazione assai atipica dopo che è trascorso dall' istante iniziale un tempo breve (cioè non astronomico) richiede la esatta inversione di tutte le velocità: se esse fossero invertite con un errore anche molto piccolo (purché non "astronomicamente piccolo") il sistema non tornerebbe indietro e anzi, probabilmente, continuerebbe ad evolvere come se nulla fosse accaduto.

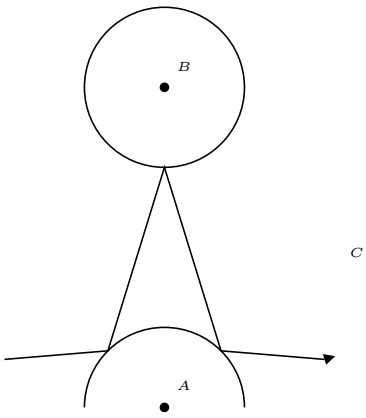
Il secondo punto di vista ha carattere più matematico e cerca di rendere quantitativo l' argomento appena descritto collegandolo all' equazione di Boltzmann.

Si immagina di considerare un dato iniziale in cui le particelle, sferette di raggio R, sono distribuite indipendentemente nello spazio delle fasi; si suppone che la densità con cui ciascuna di esse è distribuita sia $\rho f(\underline{p}, \underline{q})$ ove f è normalizzata a 1: $\int f_0(\underline{p}, \underline{q})\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{q} = 1$

Si fa evolvere questo sistema con le equazioni di Hamilton e all' istante t si suppone che sia descritto da $\rho f_t(\underline{p}, \underline{q})$, senza però supporre che le particelle siano distribuite indipendentemente; questo significa che la $\rho f_t(\underline{p}, \underline{q})$ fornisce solo l' informazione sul numero di particelle in $\mathrm{d}\underline{p}\, \mathrm{d}\underline{q}$ ma non le loro correlazioni, che saranno in generale non banali perché appunto la stosszahlansatz non sarà valida.

Si immagina ora di far tendere $\rho \to \infty$ e $R \to 0$ in modo che $\rho R^3 \to 0$ ma $\rho R^2 = l =$ quantità fissata (si considera cioè il limite di Grad-Boltzmann).

Se la discussione qualitativa presentata sopra è corretta e se si osserva che nel limite considerato il gas diviene un gas perfetto (perché le particelle divengono puntiformi) in cui l' equilibrio è raggiunto in virtù di urti fra particelle senza che mai due particelle collidano più di una volta (perché $R \to 0$ implica questo, essendo facile stimare la probabilità di ricollisione (per unità di tempo), cioè dell' evento in figura:



(7.15)

la traiettoria di C che collide due volte con A (A e B sono immaginati fissi per semplificare il disegno).

come proporzionale a ρR^3), si deve concludere che l' evoluzione del $\lim_{R\to 0} f_t = \overline{f}_t$ dovrebbe essere descritta dalla equazione di Boltzmann che, per $\overline{f}_t$, è:

$$\frac{\partial \overline{f}_t}{\partial t} - \frac{\underline{p}}{m}\cdot\frac{\partial \overline{f}_t}{\partial \underline{q}} = (\rho R^2)\int \frac{|\underline{p}' - \underline{p}''|}{m}\frac{\sigma(\underline{\omega})}{R^2}\cdot \\ \cdot (\overline{f}_t(\underline{p}', \underline{q})\overline{f}_t(\underline{p}'', \underline{q}) - f_t(\underline{p}, \underline{q}) f(\underline{p}_2, \underline{q}))\, \mathrm{d}\underline{p}_2\, \mathrm{d}\underline{\omega} \tag{7.16}$$

che ,si noti, è indipendente da ρ e R perché ρR^2 e $\sigma(\underline{\omega})/R^2$ sono indipendenti da R (perché si considera il caso delle sfere dure), $\sigma(\underline{\omega})/R^2 = 4\pi$.

Dunque l' equazione di Boltzmann descriverebbe correttamente l' evoluzione di un gas rarefatto per tempi t arbitrari: poiché ci si aspetta che nel limite di Grad Boltzmann i tempi di ricorrenza si allunghino indefinitamente mentre le collisioni facciano evolvere il sistema su una scala di tempo fissata dal valore del tempo di volo: $((\rho R^2)\overline{v})^{-1}$. Lo si vede anche dalla (7.16), in cui la scala di tempo è fissata da $\rho R^2|\underline{p}' - \underline{p}''|/m$ che "in media" è appunto $\sim \rho R^2\overline{v}$.

La (7.16) è stata dimostrata in assoluto rigore matematico solo recentemente e per tempi $t \ll 1/\rho R^2\overline{v}$ nel caso di un sistema di sfere rigide e di interessanti classi di dati iniziali f_0 (teorema di Lanford sulla congettura di Grad).

Questa è un'importante conferma, matematicamente rigorosa, del punto di vista di Boltzmann secondo il quale la reversibilità, con gli associati tempi di ricorrenza, non è in contrasto con la osservazione sperimentale della irreversibilità perché la scala di tempo su cui si manifesta la reversibilità non è osservabile mentre quella su cui si manifesta la irreversibilità è osservabile ed è legata al "tempo di volo" $(\rho R^2\overline{v})^{-1}$; inoltre la irreversibilità è compatibile con la ipotesi ergodica e l' equazione di Boltzmann fornisce un modello dello sviluppo dei moti irreversibili in situazioni in cui i tempi di ricorrenza sono enormemente più lunghi del tempo di percorrenza del libero cammino medio (ossia del tempo di volo).

Dunque il teorema di Lanford, pur avendo scarso interesse applicativo per la brevità dell' intervallo di tempo sul quale ha validità, $t \ll (\rho R^2\overline{v})^{-1}$, ha un'importanza concettuale enorme e non ancora valutata appieno da molti Fisici, perché appunto mostra in modo matematicamente preciso e rigoroso che *non c'è incompatibilità fra l' equazione di Boltzmann descrivente evoluzioni irreversibili e le equazioni reversibili di Hamilton che descrivono i dettagli del moto microscopico.*

Con queste considerazioni si esaurisce un'analisi dei fondamenti della MS classica.

Si è visto come la MS classica sia valida solo in certe condizioni (c.f.r. ad esempio la discussione del §2) almeno come formulata qui: restano da analizzare le conseguenze dell' analisi per dedurne applicazioni ed una più chiara comprensione dei suoi limiti di validità.

Tale comprensione si basa come già osservato sulle conseguenze della teoria stessa e non può essere svolta a priori come mostra ad esempio il fatto che la condizione base del §2, $\theta_+/\theta_- > 1$, è compatibile con valori della temperatura molto ragionevoli per la "Fisica di tutti i giorni" solo perché la intensità ε della energia di interazione molecolare ha ordine di grandezza di circa 10^{-14}erg: se questo dato sperimentale fosse molto più grande la condizione $\theta_+/\theta_- > 1$ potrebbe essere impossibile da soddisfare a temperature importanti per le osservazioni usualmente coperte dalla termodinamica classica, si vedano le voci "insiemi statistici" e "critica della meccanica Statistica" per una discussione di questi ultimi punti.

Insiemi statistici

per il DIZIONARIO DELLE SCIENZE FISICHE
edito dall' istituto dell' Enciclopedia Italiana
Autore della voce: Giovanni Gallavotti
versione preliminare: giugno 1984
versione 2: febbraio 1986

§1 *Gli insiemi statistici come modelli di termodinamica.*
§2 *Insiemi canonico e microcanonico: ortodicità.*
§3 *La equivalenza fra gli insiemi canonico e microcanonico.*
§4 *Non equivalenza fra gli insiemi canonico e microcanonico. Transizioni di fase. La costante di Boltzmann.*
§5 *L' insieme gran canonico ed altri insiemi statistici ortodici,*
§6 *Alcuni aspetti tecnici.*

§1 Gli insiemi statistici come modelli di termodinamica.

Nella meccanica statistica classica gli stati microscopici di un sistema di N particelle (identiche, per semplicità) di massa m, che si possono muovere in un volume V dato, si rappresentano a mezzo di cellette di ugual volume h^{3N} dello "spazio delle fasi"; le cellette hanno dimensioni δp e δq nelle coordinate di impulso e di posizione e rappresentano la massima risoluzione con cui gli stati microscopici sono osservabili: il parametro $h = \delta p \delta q$ rappresenta empiricamente la precisione con cui si possono individuare gli stati microscopici, si veda Meccanica Statistica Classica, §1.

La evoluzione temporale trasforma in un tempo τ opportunamente piccolo cellette in altre cellette: la celletta Δ è trasformata in $\Delta' = S\Delta$ da una trasformazione S definita in termini della funzione energia totale $E(\Delta)$, somma dell' energia cinetica totale $T(\underline{p})$ e dell' energia potenziale totale $\Phi(\underline{q})$:

$$E(\underline{p},\underline{q}) = T(\underline{p}) + \Phi(\underline{q}) = \sum_{i=1}^{N} p^2/2m + \sum_{i<j} \varphi(\underline{q}_i - \underline{q}_j) \qquad (1.1)$$
$$E(\underline{p},\underline{q}) \geq U^\circ = \min E(\underline{p},\underline{q}) > -\infty$$

ove $\underline{p} = (\underline{p}_1, ..., \underline{p}_N)$, $\underline{q} = (\underline{q}_1, ..., \underline{q}_N)$ sono le coordinate di impulso e posizione delle N particelle e φ è il potenziale della interazione fra le particelle (si veda Meccanica Statistica, §2). La seconda delle (1.1) è una condizione di stabilità che supporremo verificata per ogni N (con U° N–dipendente): senza di essa molti degli integrali che scriveremo sarebbero divergenti: per il suo significato fisico si veda la (2.17) e la voce "stabilità della Materia".

Si considerano poi le distribuzioni stazionarie di probabilità μ che ad ogni celletta, cioè ad ogni stato microscopico, Δ associano la sua probabilità $\mu(\Delta)$ in modo che $\mu(\Delta) = \mu(S\Delta)$.

Le famiglie $\mathcal{E}$ di distribuzioni stazionarie si possono identificare con famiglie di stati di equilibrio macroscopico in cui una generica grandezza osservabile f, cioè una generica funzione definita sulle cellette dello spazio delle fasi, assume valor medio:

$$\overline{f} = \sum_{\Delta} \mu(\Delta) f(\Delta) \qquad (1.2)$$

Data una famiglia $\mathcal{E}$ di distribuzioni stazionarie sullo spazio degli stati microscopici si possono considerare i valori medi che le osservabili più importanti assumono in uno stato $\mu \in \mathcal{E}$:

$$U(\mu) = \sum_{\Delta} \mu(\Delta) E(\Delta) \qquad \text{"energia"}$$

$$\begin{aligned} V(\mu) &= V(\Delta) = V && \text{“volume”} \\ K(\mu) &= \sum_{\Delta} \mu(\Delta) T(\Delta) && \text{“energia cinetica”} \\ P(\mu) &= \sum_{\Delta} \mu(\Delta) P(\Delta) && \text{“pressione”} \end{aligned} \tag{1.3}$$

ove $P(\Delta)$ è la variazione di impulso per unità di tempo e di superficie subita dalle particelle dello stato microscopico Δ nelle collisioni con le pareti del contenitore, cioè $P(\mu)$ è la forza per unità di superficie esercitata sulle pareti del contenitore (si veda Meccanica Statistica, §5).

Data dunque una famiglia $\mathcal{E}$ di distribuzioni stazionarie nello spazio degli stati microscopici, che chiameremo un “insieme statistico”, (a volte chiamato “ ensemble” nella terminologia di Gibbs o “ monodo” in quella di Boltzmann) si possono associare ad ogni stato $\mu \in \mathcal{E}$, “stato macroscopico”, le quantità U, V, K, P (energia, volume, energia cinetica e pressione medie) e ci si può domandare se l' insieme statistico $\mathcal{E}$ fornisca un “modello di termodinamica” in cui la temperatura assoluta si possa identificare con la energia cinetica media per particella a meno di un fattore di proporzionalità che, per semplificare formule successive, viene scritto $2/3k$:

$$T = (2/3k)\, K(\mu)/N \tag{1.4}$$

ove k è una costante da determinare empiricamente.

Il significato preciso della locuzione “fornisce un modello di termodinamica” è il seguente (si veda Meccanica Statistica, §5,6): facendo variare μ in $\mathcal{E}$ si possono seguire le variazioni di U, V, T, P; se vale la relazione:

$$(\,\mathrm{d}U + p\,\mathrm{d}V)/T = \text{differenziale esatto} \tag{1.5}$$

allora sarà possibile, integrando (1.5), definire una funzione $S(\mu)$ su $\mathcal{E}$ in modo che fra U, V, S, T, P valgano le relazioni della termodinamica Classica in cui S ha la interpretazione di “entropia”:

$$(\,\mathrm{d}U + P\,\mathrm{d}V)/T = \mathrm{d}S \tag{1.6}$$

Si associano così ad ogni stato microscopico del sistema, cioè ad ogni $\mu \in \mathcal{E}$, le quantità U, T, S, P, V ottenendo un “modello di termodinamica”: gli insiemi statistici $\mathcal{E}$ che godono della proprietà (1.6) furono chiamati brevemente da Boltzmann “ortodi” e perciò qui riferiremo a questa proprietà chiamandola “proprietà di ortodicità” di $\mathcal{E}$ (si veda Meccanica Statistica, §5,6).

La esistenza di importanti classi di insiemi statistici fu mostrata da Bolzmann che ne fornì alcuni esempi, provvedendo anche ragioni a priori per

attendersi che tali esempi dovessero non solo dare modelli di termodinamica ma addirittura "la" termodinamica del sistema in considerazione, sperimentalmente controllabile, fondando la "ipotesi ergodica" e la "equazione di Boltzmann" → (si veda meccanica statistica →, §3,5,6,7).

Dunque la teoria degli insiemi statistici pone tre questioni:
1) esistenza e descrizione di insiemi statistici ortodici;
2) equivalenza delle termodinamiche da essi descritte;
3) confronto con l' esperienza delle equazioni di stato calcolate.

In questo paragrafo si considereranno i due insiemi statistici fondamentali definiti da Boltzmann e si farà vedere la loro "ortodicità", essenzialmente secondo le idee di Boltzmann.

L' insieme canonico (si veda Meccanica Statistica Classica, §6) consiste delle distribuzioni di probabilità μ sullo spazio degli stati microscopici Δ che hanno a disposizione un volume $V = V(\Delta)$, che per semplicità supporremo cubico e con pareti perfettamente riflettenti: gli elementi μ sono parametrizzati da due parametri $\beta > 0$ e V via la relazione:

$$\mu(\Delta) = \frac{e^{-\beta E(\Delta)}}{Z(\beta, V)} \tag{1.7}$$

con

$$Z(\beta, V) = \sum_{\Delta} e^{-\beta E(\Delta)} \tag{1.8}$$

ed $E(\Delta) = E(\underline{p}, \underline{q})$, $(\underline{p}, \underline{q}) \in \Delta$, è l' energia della configurazione microscopica Δ, (1.1).

L' insieme microcanonico consiste nelle misure μ parametrizzate dai parametri U e V definite da:

$$\mu(\Delta) = \begin{cases} 1/\mathcal{N}(U,V) & \text{se } U - DE \le E(\Delta) \le U \\ 0 & \text{altrimenti} \end{cases} \tag{1.9}$$

ove $\mathcal{N}(U,V)$, chiamata "funzione di partizione microcanonica", è:

$$\mathcal{N}(U,V) = \sum_{E(D)\in(U-DE,U)} 1 = \left\{ \begin{array}{l} \text{numero delle cellette } \Delta \text{ di} \\ \text{energia } E(\Delta) \in (U - DE, U) \end{array} \right. \tag{1.10}$$

ove DE è una energia macroscopica, ma molto piccola rispetto ad U: $DE \ll U$.

In altre parole nell' insieme microcanonico si attribuisce uguale probabilità a tutte le cellette di energia macroscopica U e probabilità nulla alle altre, mentre nell' insieme canonico si dà probabilità relativa $e^{-\beta U}$ a tutte le cellette di energia macroscopica U, che però può assumere tutti i valori.

Dimostrare la ortodicità di questi insiemi statistici significa esprimere U, V, T, P in termini di due parametri (β, v), con $v = V/N$, nel caso dell'

insieme canonico, o (u, v) con $u = U/N$, $v = V/N$, nel caso dell' insieme microcanonico e, quindi, mostrare che vale la relazione (se T è $(2/3k)$ volte l' energia cinetica media per particella):

$$(\,\mathrm{d}u + p\,\mathrm{d}v)/T = \quad \text{differenziale esatto} \tag{1.11}$$

al variare di (β, v) o (u, v), rispettivamente. Si troverà che mentre l' insieme canonico è ortodico già a volume finito, l' insieme microcanonico diviene ortodico solo nel "limite termodinamico", $N \to \infty$, $U \to \infty$, $V \to \infty$ in modo che $U/N = u$, $V/N = v$ restino costanti: però è ovviamente a questo limite cui si deve essere interessati nelle applicazioni, vista la grandezza di N.

§2 Insiemi canonico e microcanonico: ortodicità

Ci sono molti altri esempi di insiemi statistici che, almeno nel limite termodinamico, sono ortodici, però prima di procedere alla discussione di altri insiemi statistici e della loro equivalenza (cioè della equivalenza dei modelli di termodinamica da essi definiti) conviene descrivere come si possa verificare la ortodicità degli insiemi canonico e microcanonico: questa verifica costituisce un punto chiave per la comprensione del pensiero di Boltzmann e dei meccanismi matematici che consentono di rendere trattabile un problema che a prima vista può apparire formidabile.

Consideriamo prima il caso dell' insieme canonico (1.7), (1.8).

La somma di partizione $Z(\beta, V)$ si può calcolare, se $h = \delta\underline{p}\ \delta\underline{q} =$ "taglia delle cellette" è piccolo, come:

$$Z(\beta, V) = \int_{\underline{q} \in V^N} e^{-\beta T(\underline{p})} e^{-\beta \Phi(\underline{q})} \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{N! h^{3N}} \tag{2.1}$$

ove il fattore $N!$ ''e introdotto per tener conto che si immagina che le N particelle del sistema siano identiche nel senso stretto del termine e, cioè, indistinguibili in linea di principio, per cui permutando le N particelle si ottengono stati microscopici descritti da cellette che devono essere considerate identiche.

Nelle (2.1) si è sostituita la somma analoga a quella che appare nelle (1.8) con un integrale; commettendo così un duplice errore:

i) un errore analitico di approssimazione dovuto al fatto che $E(\underline{p}, \underline{q}) = E(\Delta)$ solo al centro della celletta Δ;

ii) un errore combinatorio dovuto al fatto che se in una configurazione microscopica si hanno n_1 particelle in un cubetto C_1 di dimensione h^3 (pensato come un insieme nello spazio delle fasi di una singola particella), n nel cubetto C etc., allora la configurazione in questione è contata $N!/n_1!n_2!...$ volte nell' integrale (2.1) invece che $N!$ volte.

Entrambi gli errori sono ovviamente infinitesimi con h (se si intende, come immaginiamo qui, che sia δp che δq tendano a zero quando $h \to 0$) e Boltzmann li trascurò nella sua discussione.

Anche qui li trascureremo, salvo a ritornarci su *a posteriori* per valutare in quali situazioni fisiche l' ipotesi risulta ragionevole.

Anticipando il risultato di tale analisi (si veda §4), gli errori ora menzionati risulteranno trascurabili "ad alta temperatura" e, ad esempio, nel caso del gas perfetto ($\Phi(\underline{q}) = 0$), per:

$$T > T_q = (mkh^{-2}\rho^{-2/3})^{-1} \tag{2.2}$$

ove $\rho = N/V$, $k = 1.38\ 10^{-16}$ erg/$^\circ K$.

La (2.2) può essere ottenuta osservando che la rappresentazione degli stati microscopici in termini di cellette può evidentemente essere consistente solo se δp e δq sono più piccoli dei valori medi dell' impulso e della distanza fra atomi (questa condizione è meno stringente di quella esaminata al §2 della voce Meccanica Statistica (formula (2.7): $T > T_0$ che impone la compatibilità della descrizione in termini di cellette con la dinamica classica microscopica).

Poiché per la (1.4) la temperatura assoluta è tale che se $3kT/2$ è il valore medio dell' energia cinetica per particella, e coincide dunque con il valore medio di $\underline{p}_1^2/2m$, è chiaro che il valore medio dell' impulso sarà dell' ordine di $\sqrt{mkT}$ mentre il valore medio della distanza fra particelle sarà $\sqrt[3]{V/N} = \rho^{-1/3}$ e quindi $h \equiv \delta p\ \delta q < \sqrt{mkT}\rho^{-1/3}$ fornisce la (2.2). Per una discussione più dettagliata della (2.2) si veda il §4.

La (2.2) nel caso di idrogeno a densità normale, $m = 3.34\ 10^{-24} g$, $N = 2.7\ 10^{19}$ particelle in $V = 1$ cm^3, e scegliendo $h =$ costante di Planck$= 6.62\ 10^{-27}$ erg/$^\circ K$, da $T_q = 1^\circ K$.

È però bene sottolineare che se la (2.2) non è valida, e quindi non si possono trascurare le dimensioni delle cellette, viene a cadere la liceità stessa della rappresentazione degli stati microscopici in termini delle cellette e tutta la teoria dovrebbe essere ridiscussa: si vedrà infatti che in tali circostanze può divenire importante la meccanica quantistica ed in realtà la meccanica statistica classica può perder senso e validità.

La discussione del problema della ortodicità, supponendo la (2.1) corretta, senza le necessarie correzioni analitiche e combinatorie sopra esposte, equivale a scegliere $h = 0$ e, cioè si ammette la possibilità di misurare impulso e posizione di ciascuna particella simultaneamente e con infinita precisione.

Possiamo ora valutare, seguendo lo schema di Boltzmann, le grandezze termodinamiche nello stato descritto dalla distribuzione canonica di parametri β e V.

Per semplicità di notazione si identificherà la regione V occupata dal sistema con la misura V del suo volume (penseremo infatti sempre a conteni-

tori di forma cubica).

Si userà il fatto che nelle nostre approssimazioni la probabilità di trovare il sistema nello stato microscopico $\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}$ è $e^{-\beta E(\underline{p},\underline{q})}\ \mathrm{d}\underline{p}\ \mathrm{d}\underline{q}/N!h^{3N}Z(\beta,V)$, e dopo semplici calcoli si troverà:

$$\begin{aligned} K = K(\mu) &= \int \left(\sum_{i=1}^{N} \frac{\underline{p}_i^2}{2m}\right) e^{-\beta T(\underline{p})-\beta\Phi(\underline{q})} \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N}N!Z(\beta,V)} \\ v &= V/N \\ U = U(\mu) &= \frac{-\partial}{\partial\beta}\log Z(\beta,V) \\ P = P(\mu) &= \sum_{Q} \frac{N}{Z(\beta,V)} \int_{V<0} e^{-\beta T(\underline{p})} 2mv^2 \frac{s}{S} \frac{\mathrm{d}\underline{q}_2 \dots \mathrm{d}\underline{q}_N\, \mathrm{d}\underline{p}_1 \dots \mathrm{d}\underline{p}_N}{h^{3N}N!} \end{aligned} \tag{2.3}$$

ove la somma verte sui cubetti Q adiacenti alla frontiera del contenitore V via una faccia laterale di area s, mentre $S = \sum_{Q} s$ è l' area totale della superficie del contenitore cubico V e $\underline{q}_1$ è il centro di Q (si noti che $S = 8V^{2/3}$), si veda Meccanica Statistica §4,(4.4).

Con un po' di algebra facile trasformare l' ultima delle (2.3) in una forma più utile:

$$P = \beta^{-1}\frac{\partial}{\partial V}\log Z(\beta,V) \tag{2.4}$$

il calcolo è illustrato nel §6 ove sono collezionate alcune deduzioni di natura più tecnica.

A questo punto occorre solo una semplice verifica. Si definisce infatti:

$$F = -\beta^{-1}\log Z(\beta,V) \qquad \text{e} \qquad S = (U-F)/T \longleftrightarrow F = U - TS \tag{2.5}$$

e si fa uso di (2.3),(2.4) per ottenere:

$$T = (2/3k)K(\mu)/N = 1/k\beta \qquad dT/T = -d\beta/\beta \tag{2.6}$$

$$\begin{aligned} \mathrm{d}F &= (\beta^{-2}\log Z(\beta,V) + \beta^{-1}U)\,\mathrm{d}\beta - \beta^{-1}\frac{\partial}{\partial V}\log Z(\beta,V)\ \mathrm{d}V = \\ &= (F-U)\ \mathrm{d}T/T - P\ \mathrm{d}V = -S\,\mathrm{d}T - P\,\mathrm{d}V \end{aligned} \tag{2.7}$$

da cui:

$$T\,\mathrm{d}S = \mathrm{d}(F+TS) + P\,\mathrm{d}V = \mathrm{d}U + P\,\mathrm{d}V \tag{2.8}$$

che coincide con la (1.6).

Si vede anche il significato fisico della $Z(\beta,V)$: infatti la funzione $F = -\beta^{-1}\log Z(\beta,V)$ è "l' energia libera" della termodinamica.

La (2.8) dimostra l' ortodicità dell' insieme canonico.

Si noti che la (2.8) si dimostra *senza* la necessità di considerare il "limite termodinamico" $N \to \infty$, $V \to \infty$, $V/N \to v$, purché si accettino le approssimazioni che conducono alla (2.1) (ossia se h e la dimensione delle cellette si possono considerare nulle o, più fisicamente, trascurabili). Questa validità per tutti gli N e V è notevole ma si deve considerare casuale, come la discussione che segue mostra. Nel caso degli altri insiemi statistici la considerazione del limite termodinamico è necessaria per avere fra U, T, S, ρ, V le relazioni termodinamiche e, anzi, per dimostrare l' ortodocità è necessario imporre sulla energia potenziale $\Phi(\underline{q})$ descrivente le interazioni microscopiche fra le particelle, e finora piuttosto arbitraria, alcune condizioni importanti dal punto di vista fisico.

La situazione è in particolare un po' più complessa nel caso dell' insieme microcanonico perché qui è effettivamente necessario considerare il limite termodinamico. In questo caso la somma di partizione $\mathcal{N}(U, V)$

$$\mathcal{N}(U,V) = \int_J \mathrm{d}p\,\mathrm{d}q/h^{3N}N! \tag{2.9}$$

se J è l' insieme in cui si ha $(U - DE \le E(\underline{p}, \underline{q}) \le U)$, che è valida a meno degli stessi errori già discussi nel caso dell' insieme canonico.

Anche qui la ortodicità si ottiene per verifica. Definendo:

$$S = k \log \mathcal{N}(U,V) \tag{2.10}$$

e T come $(2/3k)$ volte l' energia cinetica media per particella si trova:

$$\mathrm{d}S = k\left(\frac{1}{\mathcal{N}(U,V)}\frac{\partial \mathcal{N}}{\partial U}(U,V)\ \mathrm{d}U + \frac{1}{\mathcal{N}(U,V)}\frac{\partial \mathcal{N}}{\partial V}(U,V)\ \mathrm{d}V\right) \tag{2.11}$$

e ci si chiede se il membro di destra di (2.11) possa essere scritto come $(\,\mathrm{d}U + p\ \mathrm{d}V)/T$ con P, V, T definite in (2.3).

Le derivate di $\mathcal{N}$ possono essere studiate con un po' di pazienza come nel caso dell' insieme canonico e si trova che la (2.11) può essere riscritta, cfr §6:

$$\mathrm{d}S = k(1 - 2/3N)\frac{3N}{2}\langle T(p)^{-1}\rangle\left(\ \mathrm{d}U + \frac{P\ \mathrm{d}V}{\langle T(\underline{p})\rangle^{*}\langle T(\underline{p})^{-1}\rangle}\right) \tag{2.12}$$

ove, per α reale, e se J è il dominio in cui $(U - E- \le E(\underline{p}, \underline{q}) \le U)$, si è posto:

$$\langle T(p)^{\alpha}\rangle = \frac{\int_J T(\underline{p})^{\alpha}\,\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}/h^{3N}N!}{\int_J \mathrm{d}\underline{p}\ \mathrm{d}\underline{q}/h^{3N}N!} \qquad \alpha \text{ reale} \tag{2.13}$$

$$\langle T(p)^{\alpha}\rangle = \frac{\int_{J,\underline{q}_1 \in \mathrm{d}V}\ T(\underline{p})^{\alpha}\,\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}/h^{3N}N!}{\int_{J,\underline{q}_1 \in \mathrm{d}V}\ \mathrm{d}\underline{p}\ \mathrm{d}\underline{q}/h^{3N}N!} \quad \alpha \text{ reale} \tag{2.14}$$

$\mathrm{d}V$ essendo una regione infinitesima attorno a V ottenuta spostando di η, lungo la normale esterna a V, gli elementi di superficie di V.

In altre parole $\langle T(\underline{p})^\alpha \rangle$ è il valore medio della potenza α-ma di $T(\underline{p})$, rispetto alla data distribuzione microcanonica μ, mentre $\langle T(\underline{p}) \rangle^*$ è il valore medio della energia cinetica $T(\underline{p})$ in una distribuzione μ^* ottenuta da μ imponendo che una delle N particelle si trovi vincolata a stare nella regione $\mathrm{d}V$ attorno alla superficie di V.

Se valessero le relazioni:

$$\langle T(\underline{p})^\alpha \rangle \ \langle T(\underline{p})^\alpha \rangle^* = K(\mu)^\alpha (1 + \theta_N) \tag{2.15}$$

con $\theta_N \underset{N\to\infty}{\to} 0$ e con $K(\mu)$ uguale al valore medio dell' energia cinetica nell' insieme microcanonico, si potrebbe dedurre che la (2.12) diviene, dopo averne diviso ambi i membri per N e avere fatto tendere $N \to$ mantenendo $V/N = v$ e $U/N = u$ costanti ("limite termodinamico"), la:

$$\mathrm{d}s = (du + p\, dv)/T \tag{2.16}$$

Nel caso microcanonico si vede dalle (2.12), (2.16) che la somma di partizione ha direttamente il significato fisico di entropia: $S = k \log \mathcal{N}(U, V)$. Poiché $\mathcal{N}(U, V)$ è il numero di stati microscopici di energia U e volume V, questa è la ben nota formula di Boltzmann secondo cui l' entropia è proporzionale al logaritmo del numero degli stati microscopici possibili di data energia.

Per completare l' analisi della ortodicità dell' insieme microcanonico, resta dunque da verificare le (2.15): tali relazioni possono essere dimostrate solo se si accettano, come già detto, opportune ipotesi sulla energia potenziale Φ.

Queste ipotesi, che risultano avere un significato fisico importante, sono:

a) "stabilità": esiste una costante B tale che per ogni N e per ogni configurazione $(\underline{q}_1, \ldots, \underline{q}_N) = \underline{q}$ si ha:

$$\Phi(\underline{q}) = \sum_{i<j} \varphi(\underline{q}_i - \underline{q}_j) \geq -BN \tag{2.17}$$

Questa proprietà dice che non solo l' energia potenziale è inferiormente limitata (come solitamente avviene nei sistemi meccanici realistici perché le particelle hanno un "cuore duro") ma che il suo valore minimo non può essere troppo piccolo al crescere di N.

b) "temperatezza": esistono tre costanti $C > 0$, $x > 0$, $R > 0$ per cui:

$$|\varphi(q - q')| \leq C|q - q'|^{-3-x} \text{ per } |q - q'| > R \tag{2.18}$$

Questa condizione dice che particelle molto lontane interagiscono "poco" fra loro: in base a questa ipotesi l' energia di interazione fra due blocchi

di particelle distribuite uniformemente in ciascun blocco tende a zero se la distanza fra i due blocchi viene fatta tendere all' infinito. Cioè in un sistema molto grande sottoregioni macroscopiche ma lontane interagiscono poco.

Le (2.17),(2.18) non sono verificate nel caso, speciale ma importantissimo, dei sistemi di particelle interagenti via la forza di Coulomb: il problema è costituito in realtà dalla condizione b), perché la a) è soddisfatta se si immagina che le particelle abbiano un cuore duro. La Meccanica Statistica di questi sistemi di cariche elettriche è più delicata e si veda la voce Critica della Meccanica statistica, §3, per alcuni commenti o le voci "stabilita della materia", e "plasmi" per un'analisi più dettagliata del problema.

Le (2.14) sono legate alla "legge dei grandi numeri": dicono che le variabili $T(\underline{p})$, pensate come variabili aleatorie distribuite secondo la distribuzione che corrisponde ad un elemento μ dell' insieme microcanonico o alla relativa μ^*, si veda (2.15), sono variabili a "dispersione che tende a 0" nel limite $N \to \infty$ perché il rapporto $T(\underline{p})/K(\mu)$ è tale che $\langle (T(p)/K(\mu))^\alpha \rangle$, $(T(p)/K(\mu))^\alpha \rangle^* \to 1$ per ogni α. Ossia le fluttuazioni di $T(\underline{p})^\alpha$, rispetto al suo valore medio $\langle T(p)^\alpha \rangle \simeq K(\mu)^\alpha$, non sono dell' ordine di grandezza di $\langle T(\underline{p})^\alpha \rangle$ stesso, ma molto più piccole.

Poiché $T(\underline{p})$ è somma di N variabili $\underline{p}_1^2/2m, \ldots, \underline{p}_N^2/2m$ "quasi indipendenti" ma non davvero tali (in quanto vige fra esse il vincolo $U - DE - \Phi(\underline{q}) \leq T(\underline{p}) \leq U - \Phi(\underline{q})$), è però chiaro che la (2.14) richiede una dimostrazione e non si riduce banalmente alla legge dei grandi numeri, che è formulata in termini di variabili indipendenti.

La trattazione corretta di questo problema ha dato origine ad un gran numero di altri problemi sia matematici che fisici ed è oggetto di intenso studio ormai dagli anni '60, dopo che alcuni lavori precedenti ne avevano proposto chiaramente i termini. Fino ad allora tale problema era stato poco sentito dai Fisici che consideravano soddisfacente l' argomento originale di Boltzmann.

È qui opportuno avvertire il lettore che mentre la trattazione dell' insieme canonicoqui descritta è sostanzialmente quella di Boltzmann, quella dell' insieme microcanonico è piuttosto diversa e più complessa: il motivo è che in Boltzmann vengono supposte ipotesi del tipo (2.14) solo implicitamente, e in realtà Boltzmann affronta il problema da un punto di vista leggermente diverso. Egli identifica a priori una quantità che definisce (in termini degli elementi degli insiemi statistici) come la quantità di calore dQ trasferita nel sistema in una trasformazione infinitesima e definisce ortodico un insieme statistico in cui $\mathrm{d}Q/K(\mu)$ un differenziale esatto. La trattazione dei due insiemi canonico e microcanonico è apparentemente più semplice in Boltzmann perché le definizioni di dQ nei due insiemi differiscono tra loro e, nel linguaggio che si usa qui, sono consistenti, almeno nel limite

termodinamico, solo se si suppongono valide le (2.14).

Ma non è questo il luogo per una trattazione filologicamente corretta del pensiero di Boltzmann (trattazione peraltro tuttora piuttosto insoddisfacente anche nella letteratura specializzata).

A conclusione di questo paragrafo ci si può domandare quanto l' ortodicità, degli insiemi canonico e microcanonico dipenda dall' ipotesi che la (2.1) e la (2.9) siano buone approssimazioni alle somme di partizione, o quanto la ortodicità dipenda dall' ipotesi che il sistema consista di una sola specie di particelle identiche.

Senza addentrarci nei calcoli diciamo solo che la nozione di ortodicità, nel caso che (2.1) e (2.9) vengano sostituite dalle somme esatte che approssimano, *deve essere modificata*: nel caso dell' insieme canonico si deve interpretare β come proporzionale all' inverso della temperatura assoluta mentre nel caso dell' insieme microcanonico si deve definire l' entropia direttamente via la formula di Boltzmann: $S = k \log \mathcal{N}(U, V)$. Si ottengono in tal modo due modelli di termodinamica in un senso naturale (ossia nel senso che, variando lo stato μ nell' insieme in questione e, nel primo caso, ponendo $T = 1/k\beta$ la espressione $dU + pdV)/T$ è un differenziale esatto; nel secondo ponendo $T = dS/(dU + pdV)$ la T è una grandezza ben definita, indipendente dalla trasformazione che genera le variazioni dS, dU, dV); e si potrà dimostrare che questi due modelli di termodinamica sono equivalenti. Viene però meno la identificazione universale (cioè comune a tutti gli insiemi statistici) della energia cinetica media con la temperatura assoluta che tanto ruolo ha giocato nelle fondazioni della meccanica statistica classica. Si veda la voce Equipartizione e Critica, per una' analisi dettagliata di questo punto.

Dunque gli insiemi statistici in cui le somme di partizione vengono valutate senza la "approssimazione del continuo" valida solo, come discusso, se è verificata la (2.2) possono ancora essere usati per la costruzione di modelli di termodinamica.

Tuttavia, a causa delle osservazioni seguenti la (2.2), in questi casi non è ben chiaro quale possa essere il significato della termodinamica che si costruisce: uno studio fisicamente corretto del sistema richiederebbe, in queste situazioni, l' uso della meccanica quantistica come base della trattazione.

Per quel che riguarda l' ipotesi di esistenza di un solo tipo di particelle, nei sistemi finora considerati, diciamo senza ripetere le dimostrazioni, che la ortodicità non dipende da questa ipotesi. Ci sono però alcuni cambiamenti ovvii nella formulazione e nel conteggio dei fattori combinatori. A titolo di esempio scriviamo la somma di partizione canonica per un sistema generale di N_1 particelle di tipo 1 e massa m_1, N_2 particelle di tipo 2 e massa m_2,

etc.: si ha, nell' ipotesi che la dimensione delle cellette sia trascurabile:

$$Z(\beta,V)=\frac{1}{N_1!N_2!\dots}\int\frac{\mathrm{d}\underline{p}_1\,\mathrm{d}\underline{q}_1}{h^{3N_1}}\frac{\mathrm{d}\underline{p}_2\,\mathrm{d}\underline{q}_2}{h^{3N_2}}\dots e^{-\beta\sum_\alpha T(\underline{p}_\alpha)-\beta\Phi(\underline{q}_1\dots)} \tag{2.19}$$

e la probabilità di uno stato microscopico sarà:

$$\prod_\alpha\left(\frac{\mathrm{d}\underline{p}_\alpha\,\mathrm{d}\underline{q}_\alpha}{N_\alpha!h^{3N_\alpha}}\right)e^{-\beta\sum_\alpha T(\underline{p}_\alpha)-\beta\Phi(\underline{q}_1,\dots)}Z(\beta,V)^{-1} \tag{2.20}$$

La ortodocità di questo insieme statistico, nel senso della generalizzione naturale della nozione, si mostra esattamente come nel caso di un sistema con una sola specie di particelle.

§3 L' equivalenza fra gli insiemi canonico e microcanonico.

Nei calcoli svolti nello studio degli insiemi statistici canonico e microcanonico è apparsa varie volte la costante k di Boltzmann, sempre denotata allo stesso modo ma a priori diversa caso per caso.

In realtà questa costante è una costante universale $k=1.38\ 10^{-16}$ erg/$^\circ K$.

L' iter logico che conduce alla identificazione di k ed alla dimostrazione dell' equivalenza dei modelli di termodinamica descritti dagli insiemi ortodici canonico e microcanonico è percorso in questo paragrafo.

Consideriamo dapprima il caso in cui le molecole del sistema non interagiscono, $\varphi=0$, cioè consideriamo il modello microscopico del gas perfetto.

È in questo caso facile calcolare esplicitamente le funzioni di partizione microcanonica e canonica, $\mathcal{N}$ e Z nell' approssimazione in cui la taglia delle cellette è trascurabile, si veda (2.1) e (2.9).

Si trova, eseguendo gli integrali in coordinate polari nello spazio delle $\underline{p}$:

$$\begin{aligned}\mathcal{N}(U,V)&=\frac{1}{N!h^{3N}}V^N(\sqrt{2mU}^{3N}-\sqrt{2m(U-DE)}^{3N})\underline{\omega}(3N)/3N\\ Z(\beta,V)&=\frac{V^N\sqrt{2\pi m\beta^{-1}}^{3N}}{h^{3N}N!}\end{aligned} \tag{3.1}$$

ove $\underline{\omega}(d)=\Gamma(d/2)^{-1}\sqrt{\pi}^d$ è la superficie della sfera unitaria a d dimensioni e $\Gamma(x)$ è la funzione $\Gamma\to$ diEulero.

I limiti di (3.1) per $N\to\infty$, $V\to\infty$, con $V/N=v$, $U/N=u$ fissi, si studiano facilmente via le formule di Stirling $\to\Gamma(x+1)=x^xe^{-x}\sqrt{2\pi x}(1+O(1/x))$, ovvero $N!=N^Ne^{-N}\sqrt{2\pi N}(1+O(1/N))$ e si trova, si veda il §2, (2.5),(2.10):

$$\begin{aligned}S&=k\log\mathcal{N}(U,V)=Nk(\log\frac{V}{N}+\frac{3}{2}\log\frac{U}{N}+\mathrm{cost}+O(\frac{1}{N}))\\ F&=-\beta^{-1}\log Z(\beta,V)=-N\beta^{-1}(\log\frac{N}{V}+\frac{3}{2}\log\beta+\mathrm{cost}+O(\frac{1}{N}))\end{aligned} \tag{3.2}$$

Poiché in base alla discussione del §2, S ha l' interpretazione di entropia nell' insieme microcanonico e F di energia libera $F = U - TS$ nell' insieme canonico (cfr (2.5)) si può calcolare la pressione nei due casi:

$$\begin{aligned} p/T = (\partial S/\partial V)_V = k\tfrac{N}{V}(1 + O(\tfrac{1}{N})) \quad \text{microcanonico} \\ p = -(\partial F/\partial V)_\beta = \beta^{-1} v^{-1} = kT\, v^{-1} \qquad \text{canonico} \end{aligned} \tag{3.3}$$

Se N_A è il numero di Avogadro ($N_A = 6.0\ 10^{23}$) e $N = nN_A$ (con n = numero di grammo molecole), si vede che le (2.3) stabiliscono che l' equazione di stato del gas perfetto è $PV = nRT$ in entrambi i casi, purché il valore di k si fissi uguale nei due casi e sia numericamente:

$$\begin{aligned} k = R/N = \text{ costante dei gas}/N = (8.30\ 10^7/N_A)\ \text{erg}/^\circ K = \\ = 1.38\ 10^{-16} \text{erg}/^\circ K \end{aligned} \tag{3.4}$$

Il calore specifico a volume costante risulta, dopo un facile calcolo, $3nR/2$ nei due casi, si veda la voce "Equipartizione e Critica della MS".

Come si vede, le termodinamiche previste per il gas perfetto dai due modelli microscopici canonico e microcanonico coincidono, nel limite termodinamico, e coincidono con la termodinamica sperimentalmente nota del gas perfetto, purché la costante k sia scelta nei due casi come in (3.4).

Ci si domanda ora se questa coincidenza delle termodinamiche definite da due insiemi statistici canonico e microcanonico resta vera anche per sistemi più generali.

Questo è il "problema dell' equivalenza degli insiemi statistici canonico e microcanonico". Ed è un problema di fondamentale importanza perché sarebbe assai grave per tutta la teoria se esistessero diversi insiemi statistici ortodici di distribuzioni stazionarie prevedenti per uno stesso sistema diverse termodinamiche, ossia diverse relazioni fra u, v, T, P, s tutte compatibili con le leggi della termodinamica Classica macroscopica.

Faremo vedere che si ha "in generale" equivalenza, in ogni fissato sistema, fra insieme canonico e microcanonico se la costante k che appare nella teoria di tali insiemi è la stessa.

Una volta mostrata l' equivalenza delle termodinamiche relative agli insiemi canonico e microcanonico corrispondenti ad un dato sistema ci si dovrà porre il problema se la costante k che appare come fattore di proporzionalità fra la temperatura e la energia cinetica media per grado di libertà sia la stessa per tutti gli altri sistemi, e quindi sia sempre data dalla (3.4).

Lo schema della dimostrazione della equivalenza degli insiemi canonico e microcanonico, già usato da Boltzmann e Gibbs, è il seguente.

Posto:

$$\mathcal{N}_0(U, V) = \int_{E(\underline{p},\underline{q}) \le U} \frac{d\underline{p} d\underline{q}}{h^{3N} N!} \tag{3.5}$$

si osservi che $\mathcal{N}(U,V) = \mathcal{N}_0(U,V) - \mathcal{N}_0(U-DE,V)$ e che vale la seguente relazione fra $\mathcal{N}_0$ e Z:

$$Z(\beta,V) = \beta \int_{U^0}^{+\infty} \mathrm{d}E e^{-\beta E} \mathcal{N}_0(E,V) \tag{3.6}$$

se U^0 é il minimo dell' energia e se $Z, \mathcal{N}$ sono date da (2.1),(2.9) e che si verifica integrando per parti; trattiamo qui solo il caso in cui l' approssimazione del continuo ($h \cong 0$) è lecita, (però si può vedere che l' equivalenza degli insiemi resta formalmente valida anche se non si trascurano le dimensioni delle celletteuna volta che si modifica nel senso del §2 la nozione di ortodicità). Quindi, si veda §2:

$$\begin{aligned} &F(\beta,V) = \\ &= -\beta^{-1} \log Z(\beta,V) = -\beta^{-1} \log \beta - \beta^{-1} \log \int_{U^0}^{\infty} e^{-\beta E} \mathcal{N}(E,V)\, \mathrm{d}E \end{aligned} \tag{3.7}$$

Le quantità termodinamiche specifiche (ossia per particella) nella distribuzione μ dell' insieme canonico di parametri β, V sono, al limite termodinamico ($V \to \infty, N/V = v$ fisso):

$$\begin{aligned} f_c(\beta,v) &= \lim_{N\to\infty} \frac{1}{N} F(\beta,V) && \text{energia libera canonica} \\ u_c(\beta,v) &= \lim_{N\to\infty} \frac{U(\mu)}{N} = -\frac{\partial \beta f_c}{\partial \beta}(\beta,v) && \text{energia interna canonica} \\ T_c &= \frac{1}{k\beta} = \frac{2}{3k}\frac{K(\mu)}{N}) && \text{temperatura assoluta canonica} \\ v &= \frac{V}{N} && \text{volume specifico canonico} \\ p_c &= \lim_{V\to\infty} P(\mu) = -\frac{\partial f}{\partial v}(\beta,v) && \text{pressione canonica} \\ s_c &= \frac{f_c - u_c}{T}_c && \text{entropia canonica} \end{aligned} \tag{3.8}$$

ove nell' esprimere u_c e p_c come derivate dell' energia libera f_c via le (2.3),(2.4) si è scambiata la operazione di derivazione con quella di limite senza discussione, perché si procede qui euristicamente allo scopo di esibire nella sua essenza il meccanismo di equivalenza.

Le stesse quantità calcolate nella distribuzione dell' insieme microcanonico con parametri U e V sono:

$$\begin{aligned} f_m(u_m,v_m) &= -T_m s_m + u_m && \text{energia libera microcanonica} \\ u_m &= \frac{U(\mu)}{N} = \frac{U}{N} && \text{energia interna m.c.} \end{aligned}$$

$$
\begin{aligned}
T_m &= \frac{2}{3k}\frac{K(\mu)}{N} = \frac{\partial s_m}{\partial u}(u_m, v_m)^{-1} \quad \text{temperatura assoluta m.c.} \\
v_m &= \frac{V}{N} \qquad\qquad\qquad\qquad\qquad \text{volume specifico microcanonico} \\
p_m &= P(\mu) = T_m \frac{\partial s}{\partial V^m}(u_m, v_m) \qquad\qquad \text{pressione m.c.} \\
s_m &= \lim_{N\to\infty} \frac{k}{N} \log(\mathcal{N}_0(U,V) - \mathcal{N}_0(U-DE,V)) = \qquad (3.9) \\
&= \lim_{N\to\infty} \frac{k}{N} \log \mathcal{N}_0(U,V) \qquad\qquad \text{entropia m.c.}
\end{aligned}
$$

ove le espressioni per T_m, P_m seguono dalla (2.16), quella dell' energia libera è la definizione classica e la seconda espressione per l' entropia microcanonica richiede una digressione.

Nella teoria dell' insieme microcanonico il valore di DE è lasciato non specificato (purché $DE \ll U$ e DE sia una grandezza macroscopica, cioè $DE/N \to De > 0$). Tuttavia la teoria dell' insieme microcanonico procederebbe allo stesso modo anche se $DE = U$ e si otterrebbe anche in questo caso un insieme statistico ortodico e cioè un modello di termodinamica in cui ora $S = k \log \mathcal{N}_0(U,V)$.

La funzione $\overline{s}_m(u,v) = \lim_{N\to\infty}(k/N)\log\mathcal{N}_0(U,V)$ è monotona non decrescente in u perché tale è, manifestamente, $\mathcal{N}(U,V)$, e in realtà si può dimostrare che, nei casi rilevanti, è strettamente crescente (come ci si aspetta poiché se $\overline{s}_m \equiv s_m$ allora $(\partial\overline{s}_m/\partial u)^{-1}$ sarebbe uguale alla temperatura assoluta, che dovrebbe essere positiva) almeno nei casi interessanti in cui il potenziale verifica le condizioni di stabilità e temperatezza, (2.17),(2.18), necessarie alla dimostrazione della ortodicità dell' insieme microcanonico.

Dunque, all' ordine dominante in $N \to \infty$ ed ignorando problemi di scambio di limite:

$$
\begin{aligned}
\mathcal{N}_0(U,V) &= \exp N\overline{s}_m(u,v)/k \\
\frac{\mathcal{N}_0(U-DE,V)}{\mathcal{N}_0(U,V)} &= \exp N(\overline{s}_m(u-De,V) - \overline{s}_m(u,v))/k = e^{-\alpha N} \qquad (3.10)
\end{aligned}
$$

e $\alpha > 0$ per la stretta monotonia di $\overline{s}_m$ in u e quindi i due limiti nell' ultima delle (3.9) coincidono e $s_m \equiv \overline{s}_m$. Questo dimostra anche la equivalenza delle varie versioni dell' insieme microcanonico al variare della scelta $DE = NDe$ con $De > 0$.

Tornando al problema della equivalenza fra insieme microcanonico e canonico fissiamo la costante k in (3.8), (3.9) uguale alla stessa quantità e vediamo che il problema è formulabile così: facendo corrispondere lo stato canonico di parametri $\beta = 1/kT_c$ e $v = v_c$ con quello microcanonico con parametri $u = u_m, v = v_m$ tali che $T_c = 1/k\beta = T_m$ e $v_c = v_m$ tutte le altre grandezze termodinamiche omonime devono coincidere. In tal modo,

se questa coincidenza ha effettivamente luogo, i modelli di termodinamica classica definiti dai due insiemi statistici coincideranno.

Il motivo per cui questo accade è molto semplice, se si trascurano questioni di permutabilità dei vari limiti, procedendo euristicamente. Si trova, per (3.6) e la prima di (3.10):

$$\begin{aligned} &Z(\beta, v_m) = \\ &= \beta \int_{U^0}^{\infty} e^{-\beta E} \mathcal{N}_0(E, V)\, \mathrm{d}E = N\beta \int_{U^0}^{\infty} e^{-\beta N u} e^{N S_m(u, v_m)/k}\, \mathrm{d}u = \\ &\cong N \exp N \max_u(-\beta u + \frac{1}{k} s_m(u, v_m)) \end{aligned} \tag{3.11}$$

per cui, se il massimo è raggiunto in un solo punto u_0, si deve avere che u_0 è tale che $\beta = \frac{1}{k}\frac{\partial s_m}{\partial u}(u_0, v_m)$ (perché la derivata rispetto a u si deve annullare nel punto di massimo u_0). Inoltre:

$$\begin{aligned} u_c = U(\mu)/N = \frac{\int_{U^0}^{\infty} e^{-\beta E}(E/N)\mathcal{N}_0(E, V)\, \mathrm{d}E}{\int_{U^0}^{\infty} e^{-\beta E}\mathcal{N}_0(E, V)\, \mathrm{d}E} = \\ = \frac{\int_{U^0}^{\infty} u\ e^{(-\beta u + s_m(u, v_m)/k)N}\, \mathrm{d}u}{\int_{U^0}^{\infty} e^{(-\beta u + s_m(u, v_m)/k)N}\, \mathrm{d}u} \to u_0 \end{aligned} \tag{3.12}$$

perché solo $u \cong u_0$ contribuisce agli integrali per $N \to \infty$. La (3.12) conferma anche il significato fisico di u_0 come energia interna.

In definitiva, ricordando la relazione osservata dopo le (3.11) fra u_0 e β e la $u_c = u_0$:

$$\beta = \frac{1}{k}\frac{\partial s_m}{\partial u}(u_c, v_m) = \frac{1}{k}\, T_m(u_c, v_m) \tag{3.13}$$

e scegliendo $v_c = v_m$ e u_c in modo che $T_c = T_m(u_m, v_m)$ segue che $u_c = u_m = u_0$, dalla terza delle (3.8), (3.9).

Resta da verificare che $f_m(u_m, v_m) = f_c(\beta, v_c)$; questo segue dalla (3.11) che dice anche che, per $N \to \infty$:

$$\begin{aligned} f_c(\beta, v_m) &= -\beta^{-1} \max_u(-\beta u + s_m(u, v_m)/k) = \\ &= -\beta^{-1}(-\beta u_c + s_m(u_c, v_m)/k) = \\ &= (u_c - T_c s_m(u_c, v_m)) = \\ &= (u_m - T_m s_m(u_m, v_m)) = f_m(u_m, v_m) \end{aligned} \tag{3.14}$$

perché $T_c = T_m$, $u_c = u_m$.

L' identità delle energie libere, interne e delle temperature assolute e densità implica quella delle entropie (perché gli insiemi in questione sono ortodici e quindi, fra tali grandezze, valgono le relazioni usuali della termodinamica).

§4 Non equivalenza fra gli insiemi canonico e microcanonico. Transizioni di fase. La costante di Boltzmann.

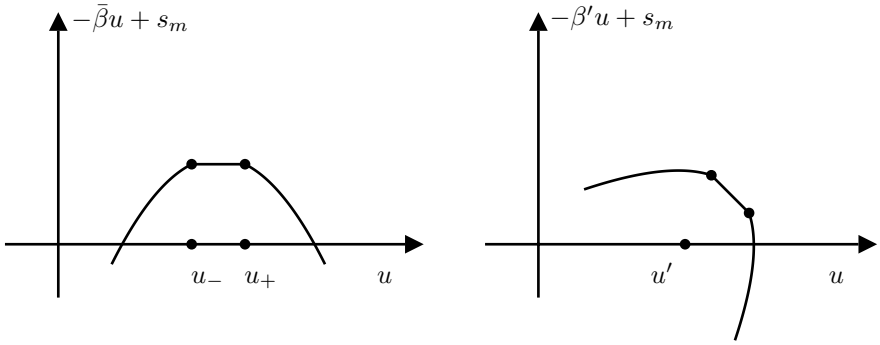
Le deduzioni del §3 sono classiche ma non rigorose: possono essere rese rigorose via un'analisi dettagliata delle proprietà qualitative delle funzioni $s_m(u,v)$ e $f_c(\beta,v)$: il punto centrale della dimostrazione sta nel far vedere che $s_m(u,v)$ è "bene approssimata" (per N grande) da $S(U,V)/N$ e inoltre è una funzione concava di u e convessa di v (mentre $f_c(\beta,v)$ è convessa in entrambe le variabili β, v). Questo implica che il massimo in (3.11) è raggiunto effettivamente in un unico punto u_0 ovvero, eventualmente, in un intervallo (u_-,u_+) ove la funzione $\beta u - s_m(u,v)$ è costante in u. La discussione dettagliata del problema, facile ma concettualmente profonda, è troppo specialistica per aver posto in questa sede e verrà pertanto tralasciata: è però opportuno menzionare che questa analisi richiede l' uso delle proprietà di stabilità e temperatezza del potenziale d'interazione φ.

Come si può prevedere dalla discussione che precede la dimostrazione della equivalenza fra insieme canonico e microcanonico non funziona più se il massimo in (3.11) è raggiunto su un intervallo (u_-,u_+), $u_- < u_+$.

Sebbene, per le proprietà generali delle funzioni concave, si può vedere che tale evento può accadere solo per valori eccezionali di β (e precisamente per un insieme di valori formanti al più un insieme numerabile), questo significa che, per valori eccezionali di β, cioè della temperatura, gli elementi corrispondenti degli insiemi canonico e microcanonico possono essere non equivalenti.

Però proprio perché tali valori di β sono eccezionali, se esistono, deve accadere che vicino quanto si vuole ad ognuno di essi, chiamiamolo $\overline{\beta}$, esistano valori β' e β'' non eccezionali ($\beta'' < \bar{\beta} < \beta'$).

Per $\beta = \beta'$ ovvero $\beta = \beta''$ si ha equivalenza fra i corrispondenti elementi degli insiemi microcanonico e canonico e in un caso l' energia interna specifica sarà $u' < u_-$ e nell' altro $u'' > u_+$ avendo denotato (u_-,u_+) l' intervallo sul quale la funzione $-g(u,v) = (-\beta u + s_m(u,v))$ assume il massimo in u per $\beta = \overline{\beta}$:



Dunque si vede che se per $\beta = \overline{\beta}$ gli stati canonico e microcanonico non sono, o possono non essere, equivalenti allora avviene che l' energia interna $u_c(\beta, v)$ subisce una discontinuità da u_- ad u_+ quando β viene fatto variare passando attraverso il valore $\overline{\beta}$. E, allora, anche l' entropia specifica $s_c(\beta, v)$ deve subire una discontinuità, perché $f_c(\beta, v) = u_c - T_c s_c$ è necessariamente continua essendo convessa, come detto sopra.

Quanto detto, anziché essere un ostacolo alla formulazione microscopica della termodinamica, lascia intravedere che la meccanica statistica possa essere il quadro teorico naturale nel quale studiare il fenomeno delle transizioni di fase: si vede infatti che alcune delle grandezze termodinamiche possono avere discontinuità in funzione di altre, esattamente del tipo empiricamente osservato nei fenomeni di transizione di fase.

Anzi i casi in cui in cui non si ha equivalenza fra elementi corrispondenti degli insiemi statistici canonico e microcanonico o, più in generale, i casi in cui esistono elementi corrispondenti ma non equivalenti in due insiemi statistici ortodici, possono essere presi come definizione dell' esistenza di una transizione di fase: ed è proprio questa la definizione oggi comunemente accettata.

In forma più chiara, da un punto di vista fisico, quello che accade in un caso di non equivalenza fra elementi di due insiemi statistici ortodici è che in generale gli stati di un insieme statistico non descrivono *tutte* le possibili fasi di equilibrio termodinamico del sistema (corrispondenti, per esempio, ad una data energia libera e temperatura). Cioè, un dato insieme statistico non è abbastanza ricco da avere fra i suoi elementi μ le distribuzioni statistiche che caratterizzano le varie fasi pure o le situazioni di coesistenza di fasi diverse (cioè i loro infiniti miscugli possibili): di solito in un dato insieme statistico $\mathcal{E}$ ci sarà una distribuzione μ che descrive uno dei possibili miscugli delle fasi coesistenti in equilibrio termico, ma mancheranno quelle che descrivono gli altri.

È proprio questo quello che avviene nei casi degli insiemi canonico e microcanonico, almeno nei pochi sistemi per cui la teoria può essere svolta fino in fondo: si veda la voce "Transizioni di Fase".

Si può dunque concludere, nel caso ora esaminato della teoria degli insiemi canonico e microcanonico, che tali insiemi statistici forniscono descrizioni equivalenti della termodinamica del sistema in corrispondenza dei valori dei parametri termodinamici cui non corrisponde alcuna transizione di fase: negli altri casi la eventuale non equivalenza non può essere considerata come un difetto della teoria ma deve essere fatta risalire al fatto che, in situazioni di non equivalenza, gli elementi dei due insiemi statistici che dovrebbero essere equivalenti in realtà non lo sono perché descrivono due fasi diverse, ma che possono coesistere fra loro in equilibrio termodinamico (o due miscugli diversi di fasi coesistenti).

E uno dei problemi più interessanti della Meccanica Statistica diviene allora proprio quello della ricerca e dello studio di situazioni di non equivalenza fra gli insiemi canonico, microcanonico o, più in generale, fra vari insiemi ortodici.

Concludiamo questo paragrafo tornando sulla questione della indipendenza della costante di Boltzmann k dal particolare sistema considerato. La discussione precedente mostra invero solo che la costante k che appare nella teoria dell' insieme canonico deve essere la stessa di quella che appare nella teoria dell' insieme microcanonico, se si vuole avere l' equivalenza fra le termodinamiche descritte dai due insiemi (a parte la questione delle transizioni di fase).

È però facile dare un argomento generale che mostri come k debba essere indipendente dal particolare sistema che si vuol studiare, e quindi debba essere dato dalla (2.4).

Invero ponendo in debole contatto meccanico due sistemi in equilibrio termico si forma un sistema composito che, nell' insieme canonico sarà descritto dalla distribuzione μ di parametri (β, v) per il primo insieme e dalla distribuzione μ' di parametri (β', v') per il secondo: si suppone per semplicità che ciascuno dei due sistemi contenga una sola specie di particelle. Il sistema composito sarà descritto dalla distribuzione prodotto $\mu \times \mu'$ perché i due sistemi sono indipendenti e la loro interazione meccanica è supposta trascurabile (è questo il significato preciso che si dà alla locuzione di "debole contatto" meccanico).

D'altra parte la distribuzione $\mu \times \mu'$ deve essere equivalente ad una opportuna distribuzione $\overline{\mu}$, di equilibrio e canonica, del sistema composito. Infatti si accetta che gli stati termodinamici siano descrivibili a mezzo di elementi di un insieme ortodico (salvo i casi eccezionali, in cui si hanno transizioni di fase, che escludiamo pensando eventualmente di cambiare di pochissimo i parametri termodinamici dei sistemi mantenendo però l' equilibrio termico). Quindi se Δ e Δ' sono due cellette rappresentanti gli stati microscopici dei due sistemi, $\overline{\mu}(\Delta \times \Delta')$ è proporzionale ad $\exp -\overline{\beta}(E(\Delta)+E(\Delta'))$, perché l' energia dello stato microscopico $(\Delta \times \Delta')$ è $E(\Delta) + E(\Delta')$, per l' ipotesi di debole interazione meccanica. Dunque:

$$\exp -\beta E(\Delta) - \beta' E(\Delta')) \qquad \exp -\overline{\beta}(E(\Delta) + E(\Delta')) \tag{4.15}$$

per ogni coppia di cellette Δ e Δ' e quindi $\beta = \beta' = \overline{\beta}$.

Ma $\beta = 1/kT$, $\beta' = 1/k'T$, $\overline{\beta} = 1/\overline{k}T$ ove T è il valore, comune per l' ipotesi di equilibrio termico, della temperatura dei tre sistemi e $k, k', \overline{k}$ sono i tre valori della costante k nei tre sistemi.

Dunque $k = k' = \overline{k}$ e cioè k è una costante universale che può essere dedotta, come già fatto, dalla teoria del gas perfetto, si veda la (3.4).

§5 L' insieme gran canonico ed altri insiemi statistici ortodici

È facile vedere che esiste un *gran numero* di altri insiemi statistici ortodici.
Ad esempio la seguente generalizzazione dell' insieme microcanonico (in cui DE viene preso uguale ad U, invece che piccolo rispetto ad U):

$$\begin{aligned} &\mu(\Delta) = 1/\mathcal{N}_0(U,V) && \text{se } E(\Delta) \le U \\ &\mu(\Delta) = 0 && \text{altrimenti} \end{aligned} \tag{5.1}$$

già considerata dopo la (2.8) è un insieme statistico ortodico per i motivi addotti al §3.
Questo insieme statistico è chiamato, anche, "microcanonico", sebbene impropriamente perchè tale nome è stato introdotto per il caso $DE = NDe$, $De > 0$, $De \ll U/N$.
Una diversa e vasta classe di insiemi statistici ortodici può essere costruita immaginando di fissare altre particelle in posizioni $\underline{q}^*_1, \underline{q}^*_2, \ldots$, e modificando $\Phi(\underline{q})$, si veda (1.1), in $\Phi^*(\underline{q})$:

$$\Phi^*(\underline{q}) = \Phi(\underline{q}) + \sum_{i=1}^{N} \sum_{j} \varphi(\underline{q}_1 - \underline{q}^*_j) \tag{5.2}$$

ove la somma su $\underline{q}^*_j$ verte sui punti $\underline{q}^*_j$ *esterni* al volume V in cui le particelle del sistema sono libere di muoversi. La energia Φ^* ha il significato di energia potenziale in presenza di particelle, fisse, all' *esterno del contenitore*. Al variare del contenitore si immagina di rimuovere le particelle fisse le cui posizioni cadono in V.
A partire dalla (5.2) si formano gli insiemi statistici microcanonico o canonico con energia $E(\underline{p},\underline{q}) = T(\underline{p}) + \Phi^*(\underline{q})$.
Se le particelle fisse vengono distribuite in modo ragionevole, ad esempio in modo che ogni cubetto unitario ne contenga un numero limitato o lentamente crescente in funzione della distanza del cubetto dall' origine (cioè se le particelle fisse vengono distribuite grosso modo uniformemente) si può dimostrare che gli insiemi statistici così ottenuti sono ortodici nel limite termodinamico ($V \to \infty, V/N = v$, $U/N = u$ fissi o $V \to \infty$, $V/N = v$, β fissi) purché ϕ verifichi le condizioni di stabilità e temperatezza del §2; se non si trascura la taglia delle cellettesi applica anche a questi insiemi lo stesso commento, posto alla fine del §1 sulla nozione di ortodicità.
Tali insiemi statistici sono denominati "insiemi microcanonico (o rispettivamente canonico) con condizioni al contorno di particelle fisse", e si può dimostrare che sono equivalenti, in assenza di transizioni di fase, all' usuale insieme canonico, nel senso analogo a quello discusso nei paragrafi precedenti: e il procedimento tecnico della deduzione è lo stesso di quello visto nel §3.

Altri insiemi statistici ortodici possono essere ottenuti facendo variare N o V, ossia considerando simultaneamente stati microscopici descriventi sistemi con diversi numeri N di particelle o occupanti diversi volumi V.

Ad esempio, assai importante nelle applicazioni è "l' insieme gran canonico" i cui elementi dipendono da due parametri $\beta > 0$ e λ reale e sono distribuzioni di probabilità, sulle cellette Δ che rappresentano gli stati di un sistema di N particelle in un volume V, fissato, con $N = 0, 1, 2, ...$, e si ha, se $E(\Delta) = E(\underline{p}, \underline{q}) = T(\underline{p}) + \Phi(\underline{q})$ e se $N(\Delta) =$ numero di particelle nello stato microscopico Δ:

$$\mu(\Delta) = \frac{e^{-\beta\lambda N(\Delta) - \beta E(\Delta)}}{\Xi(\lambda, \beta)} \tag{5.3}$$

ove il denominatore è detto "funzione di partizione gran canonica":

$$\Xi(\lambda_0 \beta) = \sum_{\Delta} e^{-\beta\lambda N(\Delta) - \beta E(\Delta)} \tag{5.4}$$

e il limite termodinamico consiste semplicemente nel far tendere $V \to \infty$ a λ, β fissi.

Più in generale si può sostituire $\Phi(\underline{q})$ con $\Phi^*(\underline{q})$ descritto in (5.2); in quest'ultimo caso si parla di "insieme gran canonico con condizioni al contorno di particelle fisse".

Una ulteriore classe di insiemi statistici ortodici è fornita dall' "insieme di pressione" e dalle sue varianti "con condizioni al contorno di particelle fisse".

In questo insieme statistico N è fisso ma il contenitore V è pensato variabile e suscettibile di assumere, mantenendo la stessa forma, vari valori del volume $V_1 = V,\ V_2 = 2V,\ V_3 = 3V, ...$ etc.

Se Δ è una celletta che descrive uno stato microscopico di N particelle pensate racchiuse in un volume $V(\Delta)$ ed avente energia $E(\Delta)$ si pone, dati $p > 0, \beta > 0$:

$$\mu(\Delta) = \frac{e^{-\beta p V(\Delta) - \beta E(\Delta)}}{J(\beta, p)} \tag{5.5}$$

ove il denominatore è detto "funzione di partizione (dell' insieme) di pressione", e

$$J(\beta, p) = \sum_{\Delta} e^{-\beta p V(\Delta) - \beta E(\Delta)} \tag{5.6}$$

Il limite termodinamico consiste semplicemente nel far tendere N all' infinito.

La teoria degli insiemi gran canonico e di pressione, come pure la teoria degli altri insiemi statistici "con condizioni al contorno di particelle fisse", può essere svolta mostrandone l' equivalenza con l' insieme statistico canonico (o microcanonico) con il metodo del massimo già visto nel caso degli

stessi insiemi canonico al §2, purché φ verifichi le proprietà di stabilità e temperatezza del §2.

A titolo di esempio di un argomento assai generale e comune in Meccanica Statistica deduciamo le proprietà dell' insieme gran canonico da quelle dell' insieme canonico mostrandone l' equivalenza, ma procedendo euristicamente (ignorando, cioè, questioni di rigore quali lo scambio di limiti).

Se μ è l' elemento generico dell' insieme gran canonico corrispondente ai parametri λ, β si ha, si veda (3.7) e la prima delle (3.8):

$$\begin{aligned}\Xi(\lambda,\beta) &= \sum_{N=0}^{\infty} e^{-\beta\lambda N} \sum_{\Delta, N(\Delta)=N} e^{-\beta E(\Delta)} = \\ &= \sum_{N=0}^{\infty} e^{-\beta\lambda N} Z_n(\beta, V) \cong \sum_{N=0}^{\infty} e^{-\beta\lambda N} e^{-\beta N f_c(\beta, V/N)} \\ &= \sum_{N=0}^{\infty} \exp V(-\beta\lambda v^{-1} - \beta v^{-1} f_c(\beta, v))\end{aligned} \tag{5.7}$$

ove nell' ultima somma $v = V/N$ e $Z_N(\beta, V)$ è la funzione di partizione canonica per N particelle nel volume V con temperatura $T = 1/k\beta$. Dunque per $V \to \infty$ e se v_0 è il valore in cui $\beta\lambda v^{-1} + \beta v^{-1} f_c(\beta, v)$ raggiunge il massimo, che supponiamo unico, troviamo:

$$\lim_{V\to\infty} (1/V) \log \Xi(\beta, \lambda) = \beta\lambda v_0^{-1} + \beta v_0^{-1} f_c(\beta, v_0) \tag{5.8}$$

e qui v_0 verifica, se si ricorda che per la (3.8) è $p_c = -v^{-1}\ \frac{\partial f_c}{\partial v}(\beta, v)$:

$$\frac{\partial}{\partial v^{-1}} (\beta\lambda v^{-1} + \beta v^{-1} f_c(\beta, v)|_{v=v_0} = 0 \to \lambda + f_c(\beta, v_0) + v_0 p_c(\beta, v_0) = 0\ . \tag{5.9}$$

D'altra parte la grandezza v_0 ha l' interpretazione di volume specifico gran canonico v_g:

$$\begin{aligned} v_g^{-1} = N(\mu)/V &= \frac{\sum_{N=0}^{\infty} \frac{N}{V} e^{-\beta\lambda N} Z_n(\beta, V)}{\sum_{N=0}^{\infty} e^{-\beta\lambda N} Z_n(\beta, V)} = \\ &= \frac{\sum_{N=0}^{\infty} N V^{-1} e^{-(\beta\lambda v^{-1} + \beta v^{-1} + \beta v^{-1} f_c(\beta, v))V}}{\sum_{N=0}^{\infty} e^{-(\beta\lambda v^{-1} + \beta v^{-1} f_c(\beta, v))V}} \xrightarrow[V\to\infty]{} v_0^{-1} \end{aligned} \tag{5.10}$$

Dunque dalla (5.9) si trova il significato fisico di λ:

$$-\lambda N = F + PV = U - TS + PV = N(f_c(\beta, v_g) + v_g P_c(\beta, v_g)) \tag{5.11}$$

ossia $-\lambda N$ è il potenziale di Gibbs canonico corrispondente ai parametri (β, v_g). E inoltre, dalla (5.8), si trova che:

$$\lim_{V\to\infty} (1/V) \log \Xi(\lambda, \beta) = p_c(\beta, v_g) \tag{5.12}$$

e cioè la funzione di partizione gran canonica è direttamente legata alla pressione canonica corrispondente ai parametri (β, v_g).

Quanto detto suggerisce che l' insieme gran canonico e quello canonico siano equivalenti se gli elementi di parametri (λ, β) e (β, v_g) con $v_g = v_0$, si veda (5.9), vengono fatti corrispondere.

La verifica consiste nel porre:

$$\begin{aligned} u_g &= \lim_{V\to\infty} \sum_{\Delta} \mu(\Delta) U(\Delta)/N(\Delta) \\ T_g &= \lim_{V\to\infty} (2/3k) \sum_{\Delta} \mu(\Delta) T(\Delta)/N(\Delta) \\ v_g &= \lim_{V\to\infty} \sum_{\Delta} \mu(\Delta) V/N(\Delta) \\ p_g &= \lim_{V\to\infty} \sum_{\Delta} \mu(\Delta) P(\Delta) \\ s_g &= (u_g - \lim \beta^{-1}(1/V) \log \Xi(\beta, \lambda))/T_g \end{aligned} \tag{5.13}$$

e nel far vedere l' identità fra le omonime grandezze calcolate nell' insieme gran canonico di parametri (λ, β) e in quello canonico di parametri (β, v_g).

Si trova, poiché $T_c(\beta, v) \equiv 1/k\beta$:

$$T_g = \lim_{V\to\infty} (2/3k) \frac{\sum_{N=0}^{\infty} e^{-\beta N} Z_N(\beta, V) T_c\left(\beta, \frac{V}{N}\right)}{\sum_{N=0}^{\infty} e^{-\beta N} Z_N(\beta, V)} \equiv 1/k\beta \tag{5.14}$$

e:

$$\begin{aligned} u_g &= \lim \frac{\sum_N e^{-\beta N} Z_n(\beta, V)\, u_c\left(\beta, \frac{V}{N}\right)}{\sum_N e^{-\beta N} Z_n(\beta, V)} = u_c(\beta, v_0) \\ p_g &= \lim \frac{\sum_N e^{-\beta\lambda N} Z_n(\beta, V)\, p_c\left(\beta, \frac{V}{N}\right)}{\sum_N e^{-\beta\lambda N} Z_n(\beta, V)} = p_c(\beta, v_g) \end{aligned} \tag{5.15}$$

per lo stesso ragionamento che in (5.10) conduce a $v_g = v_0$. Le (5.14), (5.15), (5.12) evidentemente mostrano che tutte le grandezze termodinamiche gran canoniche coincidono con le omonime canoniche.

La deduzione sopra esposta non è rigorosa perché coinvolge vari scambi di limiti e, inoltre, presuppone che $(-\lambda v^{-1} - v^{-1} f(\beta, v))$ abbia un massimo isolato ed unico (in v_0): in queste ipotesi e se φ verifica le proprietà di stabilità e temperatezza del §2, però, i problemi di rigore matematico possono essere risolti.

Dalla teoria dell' insieme canonico segue (anche se ne tralasciamo la non difficile discussione) che la funzione $-v^{-1} f(\beta, v)$ è convessa sia in β che in v^{-1} e, quindi, salvo pochi valori di λ (e precisamente ad eccezione di una famiglia al più numerabile di valori di λ) la funzione $-(\lambda v^{-1} + v^{-1} f_c(\beta, v))$ ha un unico punto di massimo.

Per λ al di fuori dell' insieme eccezionale ora menzionato si ha equivalenza totale fra la descrizione termodinamica degli stati di equilibrio del sistema in termini degli elementi degli insiemi gran canonico e canonico.

Per gli altri valori di λ (se ne esistono), invece, la funzione $-(\lambda v^{-1} + v^{-1} f(\beta, v))$ assume il massimo in un intervallo (v_-, v_+), come implicato dalle proprietà generali delle funzioni convesse: in tali casi le descrizioni dello stato in termini canonici o gran canonici possono non essere equivalenti. Ma la interpretazione di tale eventuale non equivalenza è di nuovo in termini del fenomeno delle transizioni di fase: l' inequivalenza deve essere interpretata attribuendola al fatto che gli elementi in questione degli insiemi canonico e gran canonico descrivono due stati diversi che possono coesistere in equilibrio termodinamico (e hanno cioè la stessa temperatura e pressione ma diverso volume specifico o diversa entropia etc...), si veda la discussione analoga del §3.

Uno dei risultati importanti della Meccanica Statistica contemporanea, nell' analisi del quale è non qui possibile addentrarsi, è stato quello di mostrare come, almeno in molti casi, ci sia completa equivalenza fra quelli che chiamerò "insiemi statistici ampliati": tali insiemi statistici si ottengono da un dato insieme statistico di distribuzioni stazionarie (si pensi all' insieme microcanonico o canonico per concretezza) aggiungendovi tutte le distribuzioni dello stesso insieme ma con condizioni al contorno di particelle fisse e fissate arbitrariamente.

In questi insiemi statistici più vasti potrà sì accadere che due stati preassegnati di diversi insiemi statistici ampliati che corrispondono agli stessi valori dei parametri termodinamici pressione e temperatura, ad esempio, non abbiano tutte le grandezze termodinamiche coincidenti (ad esempio abbiano diverso volume specifico), ma avverrà anche che per ogni elemento di un insieme ne esisterà uno dell' altro insieme statistico che descriverà esattamente la stessa termodinamica e cioè associerà a tutte le grandezze termodinamiche gli stessi valori e, addirittura, associerà la stessa distribuzione relativa di probabilità agli stati microscopici più probabili.

In altre parole si può dire anche che il fenomeno delle transizioni di fase è studiabile in un dato "insieme statistico ampliato" senza preoccuparsi del fatto che alcune fasi possano non essere viste: perché gli insiemi statistici ampliati contengono tutte le possibili fasi ed i loro miscugli.

Si noti che così visto il fenomeno delle transizioni di fase si manifesta come una "instabilità delle proprietà termodinamiche rispetto alle condizioni al contorno": ad esempio mantenendo la stessa temperatura e pressione ma cambiando le condizioni al contorno si possono ottenere stati termodinamici differenti; ossia cambiando le forze che agiscono sul bordo del sistema si possono ottenere cambiamenti delle grandezze termodinamiche intensive (quali il volume specifico, l' energia specifica, l' entropia specifica,... etc.)

anche se il contenitore è grandissimo.

In un certo senso questa è una ulteriore manifestazione della ricchezza della Meccanica Statistica: un fenomeno così complesso come quello delle transizioni di fase trova la sua naturale collocazione teorica e le basi per la sua analisi nella teoria degli insiemi statistici.

Come già visto nei casi sopra trattati, in un dato insieme statistico le grandezze fisiche macroscopiche o sono parametri che parametrizzano l' elemento dell' insieme statistico (come u e v nell' insieme microcanonico, β e v nell' insieme canonico ovvero β e λ nell' insieme gran canonico) oppure sono direttamente legate alla funzione di partizione (come l' entropia, l' energia libera e la pressione nei casi microcanonico, canonico e gran canonico rispettivamente) e si può mostrare, sebbene qui non lo faremo, che non dipendono dalle condizioni al contorno. Ovvero sono ottenute per differenziazione della funzione di partizione (come la temperatura nell' insieme microcanonico, l' energia interna e l' entropia nell' insieme canonico, la densità nell' insieme gran canonico,... etc.).

Dunque un modo di ricercare le transizioni di fase in modelli (ossia in sistemi ottenuti con specifiche scelte del potenziale d' interazione) fissando l' attenzione su un dato insieme statistico è di ricercare i valori dei parametri associati all' insieme statistico in questione (u, v nel caso dell' insieme microcanonico; β, v nel caso dell' insieme canonico etc.) in corrispondenza dei quali la funzione termodinamica corrispondente alla funzione di partizione (entropia nel caso microcanonico, energia libera nel caso canonico, pressione nel caso gran canonico) non è differenziabile.

Questo metodo è classico, e però ha il difetto di non fornire direttamente una descrizione microscopica degli stati di equilibrio termodinamico descriventi fasi diverse o coesistenti. Esso permette di individuare la locazione della transizione di fase in termini dei parametri termodinamici corrispondenti all' insieme statistico usato: ma non analizza le importanti particolarità fisiche delle possibili distribuzioni microscopiche che descrivono le varie fasi.

Invece lo studio della dipendenza degli stati di equilibrio come funzioni degli elementi di un insieme statistico ampliato ovviamente è potenzialmente più ricco di informazione e può condurre alla descrizione microscopica dei fenomeni delle transizioni di fase e della coesistenza delle fasi in quanto ad ogni possibile stato di equilibrio termodinamico è associato un elemento di un insieme statistico ampliato che descrive le probabilità relative delle configurazioni microscopiche che lo realizzano.

§6 Alcuni aspetti tecnici.

In questo paragrafo forniamo alcuni dettagli, per i lettori interessati ad una più approfondita comprensione, sulla deduzione di alcune identità matematiche usate nel §1, senza dimostrazione, e di alcuni altri passaggi tecnici utili. Dedurremo le equazioni (2.4),(2.12),(2.2).

1) Il passaggio dalla ultima delle (2.3) alla (2.4) merita senz'altro qualche commento.
Infatti nell' ultima delle (2.3) si può sfruttare la indipendenza degli integrali eseguiti rispetto alle variabili p da quelli rispetto alle variabili q e la simmetria della dipendenza dalle componenti di p dell' integrando.
Così si può sostituire $2mNv^2$ con mNv eliminando la condizione $v > 0$ e, poi, mNv^2 può essere sostituito da $N\underline{p}_1^2/3m$ usando la simmetria della dipendenza da $\underline{p}_1$ nelle tre componenti di $\underline{p}_1$.
Dunque si può sostituire l' integrale su $\underline{p}_1$ che in (2.3) è:

$$\int_{v>0} e^{-\beta \underline{p}_1^2/2m} 2mv^2 \, \mathrm{d}^3\underline{p}_1 \quad \text{con} \quad \int e^{-\beta \underline{p}_1^2/2m} \frac{\underline{p}_1^2}{3m} \, \mathrm{d}^3\underline{p}_1$$

e un semplice calcolo mostra che:

$$\int e^{-\beta \underline{p}_1^2/2m} \frac{\underline{p}_1^2}{3m} \, \mathrm{d}\underline{p}_1 = \beta^{-1} \int e^{-\beta \underline{p}_1^2/2m} \, \mathrm{d}\underline{p}_1 = \tag{6.1}$$

cosicché:

$$P = \sum_Q N\beta^{-1} \frac{s}{S} \frac{1}{Z(\beta, V)} \int_{\underline{q}_2, \dots, \underline{q}_N \in V^N} e^{-\beta E(\underline{p}, \underline{q})} \frac{\mathrm{d}\underline{q}_2 \dots \mathrm{d}\underline{q}_N \, \mathrm{d}\underline{p}_1 \dots \mathrm{d}\underline{p}_N}{h^{3N} N!} \tag{6.2}$$

e il punto $\underline{q}_1$, è, in ogni addendo di (6.2) localizzato in Q (che è così piccolo che non ha importanza dove $\underline{q}_1$ esattamente si trovi in Q).
Immaginiamo ora di variare V da V a $V + \mathrm{d}V$ aumentando il volume del contenitore spostando di η ogni elemento di area della sua superficie lungo la normale esterna.
Si vede che $\log Z(\beta, V)$ varia, usando $\mathrm{d}V = S\eta$, di:

$$\begin{aligned} &\mathrm{d} \log Z(\beta, V) = \qquad\qquad (6.3)\\ &= \sum_Q \frac{Ns\eta}{Z(\beta, V)} \int_{\underline{q}_2, \dots, \underline{q}_N \in V} e^{-\beta E(\underline{p}, \underline{q})} \frac{\mathrm{d}\underline{q}_2 \dots \mathrm{d}\underline{q}_N \, \mathrm{d}\underline{q}_N \, \mathrm{d}\underline{p}_1 \dots \mathrm{d}\underline{p}_N}{h^{3N} N!} \\ &= \mathrm{d}V \sum_Q \frac{N}{Z(\beta, V)} \frac{s}{S} \int_{\underline{q}_2, \dots, \underline{p}_N \in V} e^{-\beta E(\underline{p}, \underline{q})} \frac{\mathrm{d}\underline{q}_2 \dots \mathrm{d}\underline{q}_N \, \mathrm{d}\underline{p}_1 \dots \mathrm{d}\underline{p}_N}{h^{3N} N!} \end{aligned}$$

che, confrontando con (6.2), dimostra la (2.4).

2) Un'altra formula che richiede maggiori dettagli è il passaggio da (2.11) a (2.12).

Procedendo come nella deduzione di (6.2) si trova $P\,\mathrm{d}V$ a partire dalla espressione di P come media $P(\mu)$ rispetto alla misura microcanonica di parametri (u, V). Denotando con $\int^*$ l' integrale su $(\underline{p}, \underline{q})$ esteso al dominio delle $(\underline{p}, \underline{q})$ tali che $E - DE \leq E(\underline{p}, \underline{q}) \leq$ e $\underline{q}_1 \in \mathrm{d}V = \cup_Q Q$:

$$P\ \mathrm{d}V = \frac{N}{\mathcal{N}(U,V)} \int^* \frac{2}{3} \frac{\underline{p}_1^2}{2m} \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} = \frac{2}{3\mathcal{N}(U,V)} \int^* T(\underline{p}) \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} \tag{6.4}$$

avendo usato ancora nell' ultimo passaggio la simmetria di $T(p)$ in $p, \ldots, p$ e avendo scritto $\sum_Q s\eta \cdot = \int_{\underline{q}_1 \in dV} \cdot\ \mathrm{d}\underline{q}_1$ per scrivere in forma più elegante, eliminando formalmente le sommatorie su Q che appaiono naturalmente nella espressione della pressione in conformità alla sua definizione.

Al fine di connettere (6.4) alle derivate di $\mathcal{N}$ bisogna rendere più esplicita la dipendenza di $\mathcal{N}$ da U, valutando esattamente l' integrale (2.9) sulle variabili $\underline{p}$ in coordinate polari (che è un integrale elementare).

Se $\underline{\omega}(3N)$ è la superficie della sfera unitaria nello spazio a $3N$ dimensioni e se si definisce $w(U, \underline{q}) = \sqrt{2m(U - \Phi(\underline{q}))}$, si deduce:

$$\mathcal{N}(U,V) = \int \frac{\mathrm{d}\underline{q}}{h^{3N} N!} (w(U,\underline{q})^{3N} - w(U - DE, \underline{q})^{3N}) \frac{\underline{\omega}(3N)}{3N} \tag{6.5}$$

e quindi:

$$\frac{\partial \mathcal{N}}{\partial U} = \int_{V^N} \frac{\mathrm{d}\underline{q}}{h^{3N} N!} \frac{3N}{2} 2m (w(U,\underline{q})^{3N-2} - w(U - DE, \underline{q})^{3N-2}) \tag{6.6}$$

e ritornando nelle coordinate originarie:

$$\frac{1}{\mathcal{N}} \frac{\partial \mathcal{N}}{\partial U} = \frac{3N}{2} \frac{1}{\mathcal{N}} \int \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} \frac{3N-2}{3N} \frac{1}{T(\underline{p})} = \frac{3N}{2} (1 - 2/3N) \langle T(\underline{p})^{-1} \rangle \tag{6.7}$$

ove $\langle T(\underline{p})^\alpha \rangle$ è definito dalla (2.13) e l' integrale (6.7) è esteso al dominio in cui si ha $(U - DE \leq E(p,q) \leq U)$.

Procedendo invece come nella deduzione della (2.4) nel caso dell' insieme canonico, (si veda la (6.3)), si ha:

$$\begin{aligned} \frac{1}{\mathcal{N}} \frac{\partial \mathcal{N}}{\partial V}\ \mathrm{d}V &= \frac{N}{\mathcal{N}} \int^* \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} = \\ &= \frac{2N}{3} \frac{\left(\int^* T(\underline{p}) \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} / \mathcal{N} \right)}{\left(\frac{2}{3} \int^* T(\underline{p}) \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} \Big/ \int^* \frac{\mathrm{d}\underline{p}\ \mathrm{d}\underline{q}}{h^{3N} N!} \right)} \end{aligned} \tag{6.8}$$

avendo, nell' ultimo passaggio, moltiplicato e diviso per la stessa quantità, e avendo usato la notazione $\int^*$ di (6.4). Allora (6.8) e (6.4) implicano:

$$\frac{1}{\mathcal{N}}\frac{\partial \mathcal{N}}{\partial V}\,\mathrm{d}V = N\frac{P\,\mathrm{d}V}{\frac{2}{3}\langle T(\underline{p})\rangle^*} \tag{6.9}$$

ove $\langle T(\underline{p})^\alpha\rangle^*$ è definito in (2.13).

La (2.12) segue ora da (2.11), (6.7), (6.9).

3) Infine deduciamo la (2.2) nel caso semplice, lì considerato, di un gas perfetto, $\varphi = 0$.

Se si immagina di dividere lo spazio delle fasi a 6 dimensioni descrivente delle singole particelle del sistema, in cellette C aventi la forma:

$$C = \quad \text{insieme delle } (\underline{p},\underline{q}) \quad \text{in } R^3 \quad \text{tali che:}$$

$$\begin{cases} k'_\alpha\delta q - \delta q/2 \le q_\alpha \le k'_\alpha\delta q + \delta q/2 \\ k_\alpha\delta p - \delta p/2 \le p_\alpha \le k_\alpha\delta p + \delta p/2 \end{cases} \qquad \alpha = 1,2,3 \tag{6.10}$$

e $\underline{k},\underline{k}'$ sono due vettori a componenti intere, si vede che l' energia di una particella che si trova in C è $\varepsilon(C)$:

$$\varepsilon(C) = \sum_{\alpha=1}^{3} \underline{k}_\alpha^2\delta p^2/2m \tag{6.11}$$

Inoltre uno stato microscopico Δ del sistema può essere assegnato dando i "numeri di occupazione" n_C, per ogni celletta, che dicono quante particelle occupano una data celletta. Allora, senza errori combinatorii né analitici:

$$Z(\beta,V) = \sum_{\substack{n_G\ge 0 \\ \sum_C n_C = N}} e^{-\beta\sum_C n_C\varepsilon(C)} \tag{6.12}$$

invece della espressione *affetta da errori combinatorii* evidenti (in cui si denota $L = V^{1/3}$ il lato del contenitore):

$$Z(\beta,V) = \sum_{\substack{n_C\ge 0 \\ \sum_C n_C = N}} \frac{1}{(\prod_C n_C!)} e^{-\beta\sum_C n_C\varepsilon(C)} = \frac{1}{N!}\left(\sum_C e^{-\beta\varepsilon(C)}\right)^N =$$

$$= \tfrac{1}{N!}\left(\tfrac{L}{\delta p}\textstyle\sum_{k=-\infty}^{+\infty} e^{-\beta k^2(\delta p)^2/2m}\right)^{3N} = \tfrac{1}{N!}\left(\tfrac{L\delta q}{h}\textstyle\sum_{-\infty}^{+\infty} e^{-\beta k^2(\delta p)^2/2m}\right)^{3N} \tag{6.13}$$

$$= \tfrac{1}{N!}V^N\sqrt{2m\beta^{-1}}^{3N}\left(\tfrac{1}{\sqrt{2m\beta^{-1}}}\delta p\textstyle\sum_{k=-\infty}^{+\infty} e^{-\beta k^2(\delta p)^2/2m}\right)^{3N}$$

che poi conduce alla (2.1) se $h \cong 0$ (ossia alla seconda delle (3.1), perchè $\varphi = 0$) e se si approssima la somma nell' ultimo membro di (6.13) con

il corrispondente integrale $\sqrt{\frac{\beta}{2m}} \int e^{-\beta \underline{p}^2/2m}\, \mathrm{d}p$, commettendo così anche l' errore analitico sopra discusso.

Per confrontare (6.12) con (6.13) o (2.1) è chiaro che si deve decidere se i valori di n_C che danno il principale contributo alla (6.12) sono solo quelli per cui $n_C = 0, 1$ (caso in cui (6.12) e (6.13) sono buone approssimazioni l' una dell' altra e di (2.1) perché il fattore $n_C!$ vale 1 nella maggior parte dei casi importanti).

Si deve quindi calcolare il valore medio $\overline{n_C}$ della grandezza n_C rispetto alla distribuzione canonica e la condizione di trascurabilità dell' errore combinatorio sarà $\overline{n_C} \ll 1$.

Nell' insieme canonico, per la definizione (1.7), la probabilità di trovare una particella, della quale è data la posizione, con un impulso in $\mathrm{d}\underline{p}$ è ("legge di Maxwell–Boltzmann"):

$$(\exp -\beta \underline{p}^2/2m)\, \mathrm{d}\underline{p}/(\sqrt{2\pi m\beta^{-1}})^3 \tag{6.14}$$

e quindi se $\rho = N/V$ è la densità del sistema si avrà:

$$\overline{n}_C = \rho(\delta q)^3 e^{-\beta \underline{p}^2/2m} \frac{(\delta p)^3}{\sqrt{2\pi m\beta^{-1}}^3} \leq \frac{\rho h^3}{\sqrt{2\pi m\beta^{-1}}^3} \tag{6.15}$$

e quindi $\overline{n}_C \ll 1$, per tutte le cellette C, se $T > T_q$ con T_q dato da (2.2).

Per quel che riguarda l' errore che abbiamo chiamato analitico è chiaro che esso sarà trascurabile se $\sqrt{\beta/2m}\delta p \ll 1$: nel nostro contesto non abbiamo fissato δp e δq separatamente; tuttavia δq deve essere scelto certamente tale che $\delta q > \rho^{-1/3}$ = distanza media fra le particelle, altrimenti non avrebbe senso parlare di particelle in quanto oggetti separatamente definiti nel sistema: così, con questa scelta di δq, da $\delta p\ \delta q = h$ si ricava $\delta p = h/\rho^{-1/3}$ e si vede che la condizione $\sqrt{\beta/2m}\delta p \ll 1$ è la stessa della (2.2.)

Equipartizione e critica della Meccanica Statistica Classica.

per il DIZIONARIO DELLE SCIENZE FISICHE
edito dall' Istituto dell' Enciclopedia Italiana.
Autore della voce: Giovanni Gallavotti.
versione preliminare: giugno 1984.
versione 2: febbraio 1986.

§1 *Equipartizione ed altri paradossi e applicazioni della Meccanica Statistica Classica.*
§2 *I problemi della Meccanica Statistica Classica quando h non è trascurabile.*
§3 *Meccanica Statistica Quantistica. Introduzione.*
§4 *Problemi della Meccanica Statistica contemporanea.*

§1 Equipartizione ed altri paradossi ed applicazioni della Meccanica Statistica

Una delle più note conseguenze della meccanica statistica classica è il " principio di equipartizione dell' energia": meno noto è che tale principio, dopo qualche abbagliante successo iniziale, si riveli proprio come il segno della inadeguatezza ed inefficacia della meccanica statistica classica stessa riguardo alla soluzione di importanti problemi che pur ricadono nel suo ambito. Così anche altre ben note, ed importanti, applicazioni sono accompagnate da gravi paradossi e problemi teorici.

Illustreremo queste affermazioni attraverso alcuni significativi esempi.

I) Il calore specifico del gas perfetto.

Usando l' insieme canonico ed immaginando che h sia molto piccolo, (si veda la voce Meccanica Statistica Classica), si calcola facilmente l' energia interna per un modello generale in cui ogni particella ha l gradi di libertà e non interagisce con le altre. Gli l gradi di libertà descrivono i tre gradi di libertà traslazionali del baricentro e gli $(l-3)$ gradi di libertà interni, che descrivono il moto interno della molecola.

Si suppone che l' energia sia una forma quadratica negli l impulsi coniugati $p_1, p_2, \ldots, p_l$ e, eventualmente, in alcune delle coordinate di posizione interne:

$$E(p,q) = \sum_{j=1}^{3} \frac{\underline{p}_j^2}{2m} + \sum_{j=4}^{l_0} \frac{p_j^2}{2M_j(\underline{\hat{q}})} + \sum_{j=l_0+1}^{l} \frac{p_j^2 + \omega_j^2 q_j^2}{2} \qquad (1.1)$$

ove $p_1, p_2, p_3, q_1, q_2, q_3$ sono le coordinate di impulso e posizione descriventi il baricentro della molecola, m ne è nella massa totale mentre $p_4, \ldots, q_4, \ldots$ sono le coordinate di impulso e posizione descriventi i gradi di libertà interni, e $\underline{\hat{q}} \equiv (\underline{q}_4, \ldots, \underline{q}_{l_0})$. La (1.1) è la forma che ci si attende per una molecola che abbia alcuni gradi di libertà interni, in numero di $l - l_0$, cui corrispondono piccole oscillazioni attorno a posizioni di equilibrio (e corrispondono ai valori di $j = l_0 + 1, \ldots, l$, ai quali sono associate le relative "frequenze proprie" $2\pi/\omega_j$).

Chiameremo, in riferimento alla (1.1), gradi di libertà traslazionali i primi l_0 gradi di libertà ed "oscillatorii" gli ultimi $l - l_0$: la variabile coniugata ad un grado di libertà traslatorio potrà essere sia una coordinata di posizione che varia in $V (j = 1, 2, 3)$ o anche una variabile angolare interna, mentre di regola la variabile q_j coniugata ad un impulso p_j di un grado di libertà oscillatorio sarà sempre una variabile (che varia arbitrariamente su tutta la retta reale) che descrive un grado di libertà interno.

Ad esempio se il gas è costituito da atomi puntiformi di massa m si ha che $l = 3$ e $E(\underline{p}, \underline{q}) = T(\underline{p}) = (p_1^2 + p_2^2 + p_3^2)/2m$.

Se il gas è costituito da molecole biatomiche fatte di due atomi a distanza fissa ρ, l' energia cinetica è:

$$T(\underline{p}) = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + \frac{1}{2\mu}\left(\frac{p_4^2}{\rho^2} + \frac{p_5^2}{\rho^2 \sin^2\theta}\right) \tag{1.2}$$

ove m è la massa totale e μ la massa ridotta del sistema ($m = m_1 + m_2, \mu = m_1 m_2/m$) e p_4, p_5 sono i momenti coniugati alle variabili, ϑ e φ, di latitudine e di azimuth della congiungente i due atomi. In questo caso le variabili q coniugate ai primi tre impulsi sono variabili in V, mentre le altre due variabili sono variabili angolari.

Per questi gas perfetti:

$$Z(\beta, V) = \int \frac{\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}}{h^{3N}N!} e^{-\beta \sum_{i=1}^{N} T(\underline{p}_i, \underline{q}_i)} \tag{1.3}$$

e quindi il valore medio dell' energia si calcola usando la fattorizzazione dei vari integrali e calcolando esplicitamente prima quelli (gaussiani) sulle $\underline{p}$ e sulle $\underline{q}$ dei gradi di libertà oscillatori e quindi quelli sulle altre coordinate q (che risultano banali se eseguiti dopo quelli sulle $\underline{p}$).

Se $\hat{\underline{q}}$ sono le coordinate di posizione traslatorie e $\tilde{\underline{q}}$ quelle oscillatorie si ha:

$$\begin{aligned} U &= \frac{\int \left(\sum_{i=1}^{N} T(\underline{p}_i, \hat{\underline{q}}_i)\right) e^{-\beta \sum_{i=1}^{N} T(\underline{p}_i, \hat{\underline{q}}_i)} \prod_i \mathrm{d}\underline{p}_i\,\mathrm{d}\underline{q}_i}{\int e^{-\beta \sum T(\underline{p}_i, \hat{\underline{q}}_i)} \prod_i \mathrm{d}\underline{p}_i\,\mathrm{d}\underline{q}_i} = \\ &= \frac{N \int \mathrm{d}\hat{\underline{q}} (\int T(\underline{p}, \hat{\underline{q}}) e^{-\beta T(\underline{p}, \hat{\underline{q}})}\,\mathrm{d}\underline{p}\,\mathrm{d}\tilde{\underline{q}})}{\int \mathrm{d}\hat{\underline{q}} (\int e^{-\beta T(\underline{p}, \hat{\underline{q}})}\,\mathrm{d}\underline{p}\,\mathrm{d}\tilde{\underline{q}})} = \\ &= \frac{N\beta^{-1} \int \mathrm{d}\hat{\underline{q}} \left(\frac{l_0}{2} + (l - l_0)\right)}{\int \mathrm{d}\hat{\underline{q}}} = N\beta^{-1}(l_0/2 + (l - l_0)) \end{aligned} \tag{1.4}$$

Questa formula è interessante perché affatto indipendente dalla forma della (1.1) (e cioè dei coefficienti $M_\alpha(\hat{\underline{q}}), \omega_i, m$), e dice che l' energia interna di un gas perfetto è data dal numero di gradi di libertà moltiplicati per $1/2\beta = kT/2$ (" equipartizione dell' energia fra i vari gradi di libertà e fra energia cinetica ed energia potenziale"), contando per due i gradi di libertà oscillatorii, perché questi ultimi contribuiscono anche all' energia potenziale. Si suol dire che c'è "equipartizione" della energia fra i vari gradi di libertà e fra energia cinetica ed energia potenziale elastica.

Allora il calore specifico di un gas monoatomico o di un gas di molecole biatomiche rigide (cioè con atomi a distanza fissa) è rispettivamente:

$$C_V = \frac{\partial U}{\partial T} = \frac{3}{2}nR \quad \text{ovvero } \frac{5}{2}nR \tag{1.5}$$

ove $n = N/N_A$, con $N_A =$, è numero di Avogadro, il numero di moli del sistema, e $R = kN_A$ è la costante dei gas.

La (1.5) è in buon accordo con l' esperienza per i gas monoatomici, meno per quelli biatomici.

In realtà la (1.5) non è accettabile in generale, neppure per i gas monoatomici, perché è noto che alcuni gas sono costituiti da atomi con molti gradi di libertà mentre il loro calore specifico è $3nR/2$, (ad esempio il neon potrebbe essere pensato come fatto di 20 protoni e neutroni e 10 elettroni, cioè avrebbe 90 gradi di libertà, dei quali 87 oscillatorii (!)).

Ma anche il caso semplice di una molecola biatomica in cui si trascurano tutti i gradi di libertà interni tranne i tre che descrivono la posizione relativa dei due atomi è concettualmente poco chiaro: se si facesse l' ipotesi di rigidità della distanza fra i due atomi il calore specifico sarebbe $5nR/2$; se invece si ammettesse che la distanza fra i due atomi oscilli elasticamente attorno ad un valore di equilibrio (il che è più realistico), il calore specifico diverrebbe $7nR/2$ perché i gradi di libertà sarebbero 6 di cui uno oscillatorio.

Dunque è chiaro che *le cose vanno come se* certi gradi di libertà interni fossero meno importanti di altri e non contribuissero all' equipartizione dell' energia che, dunque, non sarebbe in generale valida a dispetto della sua semplicissima deducibilità dalla teoria dell' insieme canonico.

II) Il calore specifico dei solidi.

Un altro successo–fallimento della MS classica è la teoria del calore specifico dei solidi cristallini.

Un solido cristallino viene schematizzato come un sistema di particelle che oscillano attorno a posizioni ideali di equilibrio disposte su un reticolo ideale, ad esempio quadrato e di passo a.

È noto dalla teoria dei solidi che un tale sistema è descritto in opportune coordinate (" coordinate normali") dall' hamiltoniana:

$$H = \sum_{\underline{k}} \frac{p_{\underline{k}}^2 + \omega(\underline{k})^2 q_{\underline{k}}^2}{2} \qquad (p_{\underline{k}}, q_{\underline{k}}) \in R^6 \tag{1.6}$$

ove la somma verte sulle terne $\underline{k} = (k_1, k_2, k_3)$ di numeri interi con $k = 0, 1, \ldots, \sqrt[3]{N} - 1$ se N è il numero di atomi del cristallo, supposto cubico e di lato a $\sqrt[3]{N} = L$, e:

$$\omega(\underline{k})^2 = 2c^2 \sum_{i=1}^{3} \left(1 - \cos \frac{ak_i}{L} 2\pi\right) \tag{1.7}$$

e c è la velocità di propagazione del suono nel cristallo.

Se si potesse descrivere il sistema in termini dell' insieme canonico si potrebbe calcolare l' energia interna come:

$$U = \frac{\frac{1}{2}\int \sum_{\underline{k}} (\underline{p}_{\underline{k}}^2 + \omega(\underline{k})^2 \underline{q}_{\underline{k}}^2)\, e^{-\beta H(\underline{p},\underline{q})}\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{q}}{\int e^{-\beta H(\underline{p},\underline{q})}\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{q}} = 6N/2\beta \tag{1.8}$$

perché la somma su k verte su $3N$ valori ed il calcolo procede come nel caso della discussione del principio di equipartizione con la differenza che ora tutti i gradi di libertà sono "oscillatorii".

Dunque il calore specifico di un cristallo dovrebbe essere:

$$C = 3Nk = 3nR \tag{1.9}$$

se n è il numero di moli, ("legge di Dulong–Petit").

Se si tiene conto, però, che un modello tipico di un solido conduttore consistere in N ioni, su un reticolo, ed N elettroni, liberi formanti un gas perfetto di elettroni, si trova invece che il calore specifico dovrebbe essere $3nR + 3nR/2$.

Sperimentalmente il calore specifico dei cristalli a temperatura alta è invero circa $3nR$ ("legge di Dulong–Petit"). A bassa temperatura invece il calore specifico tende a zero con T ("terza legge della termodinamica").

Dunque anche in un solido cristallino la meccanica statistica classica produce previsioni errate: certi gradi di libertà sono essenzialmente "congelati" perché non contribuiscono al calore specifico (il loro contributo all' energia interna è quello che darebbero in base al principio di equipartizione se la loro temperatura si potesse considerare nulla: il che è però impossibile perché allora il sistema non sarebbe in equilibrio termico). Inoltre a bassa temperatura anche i gradi di libertà descriventi le oscillazioni fra gli atomi sembrano divenire via via meno descrivibili dalla meccanica statistica classica perché il calore specifico devia dalla legge di Dulong e Petit e tende a zero.

III) Il corpo nero.

Anche la teoria della termodinamica della radiazione può essere svolta sulla base della teoria degli insiemi statistici, e anche in essa si giunge a contraddizioni stridenti con le osservazioni, se si usa la meccanica statistica classica.

Anzi fu proprio la teoria del corpo nero il luogo ove, storicamente, la contraddizione fu più sentita e condusse alle origini della meccanica quantistica e della meccanica statistica quantistica.

Se si considera una regione V, cubica, riempita di radiazione in equilibrio termico con l' ambiente e se si descrive il campo elettromagnetico tramite il

potenziale vettore $\underline{A}$ e le relazioni (conseguenza delle equazioni di Maxwell nel vuoto):

$$\underline{E} = -\frac{1}{c}\frac{\partial \underline{A}}{\partial t}\,, \quad \underline{H} = \text{rot } \underline{A}\,, \qquad \text{div } \underline{A} = 0 \tag{1.10}$$

ove c è la velocità della luce, $c = 2.99\ 10^{10}$ cm/sec, è ben noto che il moto di tale campo è descritto dalla lagrangiana:

$$\mathcal{L} = \frac{1}{8\pi}\int_V (\underline{E}^2 - \underline{H}^2)\,\mathrm{d}\underline{x} \tag{1.11}$$

Se L è il lato del volume V occupato dalla radiazione, che per semplicità conviene considerare con condizioni periodiche al contorno (identificando i lati opposti di V), si potrà scrivere A in termini della sua serie di Fourier:

$$\underline{A}(\underline{x}) = \frac{1}{\sqrt[3]{L^3}}\sum_{\underline{k}}\sum_{\alpha=1}^{2} A^{(\alpha)}(\underline{k})\underline{e}^{(\alpha)}(\underline{k})e^{i\underline{kx}} \tag{1.12}$$

ove $\underline{k} = \frac{2\pi}{L}\underline{\nu}$ e $\underline{\nu}$ un vettore a coordinate intere, e $\underline{e}^{(\alpha)}(\underline{k})$ sono due "vettori di polarizzazione" unitari ortogonali a $\underline{k}$ e fra loro.

Si trova allora:

$$\mathcal{L} = \frac{1}{2}\sum_{\underline{k}}\sum_{\alpha=1}^{2}\left(\frac{1}{4\pi c^2}\dot{A}^{(\alpha)}(\underline{k})^2 - \frac{1}{4\pi}\underline{k}^2 A^{(\alpha)}(\underline{k})^2\right) \tag{1.13}$$

Dunque il campo nella cavità è descritto dalla hamiltoniana:

$$\begin{aligned} H &= \frac{1}{2}\sum_{\underline{k}}\sum_{\alpha=1}^{2}\left(4\pi c^2 p^{(\alpha)}(\underline{k})^2 + \frac{\underline{k}^2}{4\pi}q^{(\alpha)}(\underline{k})^2\right) = \\ &= \frac{1}{2}\sum_{\underline{k}}\sum_{\alpha=1}^{2}(\overline{p}^{(\alpha)}(\underline{k})^2 + \underline{k}^2 c^2 \underline{q}^{(\alpha)}(k)^2) \end{aligned} \tag{1.14}$$

ove le coordinate $(p^{(\alpha)}(\underline{k}),\, q^{(\alpha)}(\underline{k}))$ ovvero le coordinate $(\overline{p}^{(\alpha)}(\underline{k}), \overline{q}^{(\alpha)}(\underline{k})) = (\sqrt{4c^2\pi}p^{(\alpha)}(\underline{k}),\ q^{(\alpha)}(k)/\sqrt{4c^2\pi})$ sono coordinate canonicamente coniugate (equivalenti perché la trasformazione $(\underline{p}, \underline{q}) \Leftrightarrow (\overline{\underline{p}}, \overline{\underline{q}})$ è canonica).

Dunque un campo elettromagnetico in una cavità V può essere pensato come un sistema di infiniti oscillatori armonici indipendenti.

È molto tentante descrivere questo sistema tramite la meccanica statistica classica, dicendo che alla temperatura T il sistema sarà descritto dall' insieme canonico e, quindi, la probabilità di trovare l' oscillatore di indici $(\alpha, \underline{k})$, di polarizzazione α e bettore d' onda $\vec{k}$, nella celletta $C = dp^{(\alpha)}(\vec{k})dq^{(\alpha)}(\vec{k})$ è:

$$e^{-\frac{\beta}{2}(p^{(\alpha)}(\underline{k})^2 + c^2\underline{k}^2 q^{(\alpha)}(\underline{k})^2}\,\frac{dp^{(\alpha)}(\vec{k})dq^{(\alpha)}(\vec{k})}{\sqrt{4\pi^2\beta^{-1}\underline{k}^2 c^2}}\,. \tag{1.15}$$

È chiaro che assumendo la (1.15) si suppone di trascurare le dimensioni della celletta: questo di solito introduce due tipi di errori, combinatorii ed analitici, discussi al §1, 4 della voce Insiemi Statistici. In questo caso l' errore combinatorio è assente perché questa volta gli oscillatori sono due a due distinti. Però se β è grande si fa ugualmente sentire l' errore dovuto ad aver trascurato la taglia h delle cellette, considerando $p^{(\alpha)}(\underline{k})$ e $q^{(\alpha)}(\underline{k})$ come variabili continue.

Accettando la (1.15) la energia media per oscillatore sarà, per il solito argomento di equipartizione, kT perché ogni oscillatore rappresenta un grado di libertà oscillatorio.

Dunque è chiaro che, se $\nu = |\underline{k}|c/2\pi$ è la frequenza dell' onda di numero d' onda $\underline{k}$, la quantità di energia $L^3 u_\nu d\nu$ corrispondente agli oscillatori con frequenza fra ν e $\nu + d\nu$, è legata al numero di vettori interi $\underline{n}$ tali che $\nu \leq |\underline{n}|c/L \leq \nu + d\nu$ via la relazione:

$$\begin{aligned} L^3 u_\nu d\nu &= \beta^{-1} 2(\text{numero di } |\underline{n}| \text{ tali che } |\underline{n}|c/L \in (\nu, \nu + d\nu)) = \\ &= 2\beta^{-1}(L\nu/c)^2 \frac{4\pi L}{c}\, d\nu = L^3\, \frac{8\pi}{c^3}\, \beta^{-1}\nu^2\, \mathrm{d}\nu \end{aligned} \tag{1.16}$$

ove il fattore 2 dopo β^{-1} nel primo passaggio è presente perché, per ogni $\underline{k}$, ci sono due oscillatori con polarizzazioni diverse e con la stessa energia media kT.

Dunque si trova la formula di Rayleigh–Jeans:

$$u_\nu = \frac{8\pi\nu^2}{c^3}\, kT \tag{1.17}$$

che è in manifesto contrasto con l' esperienza perché $\int_0^\infty u_\nu\ \mathrm{d}\nu = \infty$ e una cavità radiante avrebbe energia infinita.

Sperimentalmente si osserva la (1.17) se ν è piccolo e, di nuovo, per ν grandi le osservazioni sono in contrasto con il teorema di equipartizione dell' energia, perché si osserva che u_ν tende a zero rapidamente per ν che tende all' infinito.

Dunque la meccanica statistica classica, nelle tre applicazioni discusse sopra conduce a paradossi e previsioni errate.

Nel prossimo paragrafo vedremo che tali paradossi sparisano, se si tiene conto, pur sempre nell' ambito della meccanica statistica classica, che $h \neq 0$: e si potrà dire che tutte le contraddizioni e inconsistenze che si incontrano nella meccanica statistica classica appaiono quando, per semplificare le formule, si commettono gli errori analitici e combinatori equivalenti ad assumere $h \cong 0$, si veda §1, §2 e §6 della voce insiemi statistici).

§2 I problemi della meccanica statistica classica quando h non è trascurabile

Nel §1 si è sempre trascurata la grandezza h delle cellette dello spazio delle fasi rappresentanti gli stati microscopici del sistema. Come più volte sottolineato, così facendo si commettono importanti errori (si veda §1 e §4 della voce Insiemi Statistici) che in ultima analisi sono connessi ai paradossi del paragrafo precedente.

L' errore principale sta però nel fatto che, non potendosi più assimilare ogni celletta dello spazio delle fasi ad un punto, si dovrebbe in realtà rinunciare alla descrizione della statistica dei sistemi con i metodi e i concetti della meccanica classica: la rivelabilità delle minime dimensioni delle cellette implica necessariamente la riformulazione dei principi di base della meccanica e quindi della meccanica statistica.

Per rendersi conto di quanto drastici possano essere i cambiamenti della termodinamica di un sistema nel regime "quantistico" in cui h non è trascurabile, si può continuare ad assumere come valida la descrizione del sistema in termini di cellette e valutare meglio le somme di partizione dei vari insiemi statistici, evitando gli errori (combinatori ed analitici descritti nei §1,2 e §6 della voce Insiemi Statistici) trascurabili solo per $h \to 0$.

Si consideri, come primo esempio, un gas perfetto di particelle identiche prive, per semplicità, di gradi di libertà interni, e sia C una celletta generica dello spazio a 6 dimensioni nel quale si descrive lo stato di una singola particella: il volume di C sia $(\delta p\, \delta q)^3 = h^3$.

Poiché le particelle sono indistinguibili, le configurazioni microscopiche Δ sono individuate dai numeri di particelle n_C, della configurazione Δ, che occupano la celletta C. E si ha, allora:

$$\begin{aligned} E(\Delta) &= \textstyle\sum_C n_C e(C) \qquad \text{energia totale} \\ N(\Delta) &= \textstyle\sum_C n_C \qquad\quad \text{numero di particelle} \end{aligned} \tag{2.1}$$

ove $e(C)$ è l' energia di una particella nella celletta C.

Studiamo il sistema nell' insieme gran canonico, ove i calcoli sono un po' più semplici. La funzione di partizione è allora:

$$\Xi(\beta, \lambda) = \sum_{\{n_C\}} e^{-\beta\lambda \sum_C n_C} e^{-\beta \sum_C n_C e(C)} \tag{2.2}$$

ove, per ogni $C, n_C = 1, 2, 3, \ldots$: si veda Insiemi Statistici , §5 (5.4).

Eseguendo le sommatorie esplicitamente, evitando così di commettere gli errori di cui stiamo studiando gli effetti, si trova:

$$\Xi(\beta, \lambda) = \prod_C \frac{1}{(1 - e^{-\beta\lambda - \beta e(C)})} = \exp - \sum_C \log(1 - e^{-(\beta\lambda + \beta e(C))} \tag{2.3}$$

e la probabilità che $n_C = n$ si calcola subito dalla interpretazione fisica degli elementi dell' insieme gran canonico, si veda la voce Insiemi Statistici, §5:

$$p(n;C) = \frac{e^{-\beta\lambda n - \beta n e(C)}}{(1 - e^{-\beta\lambda - \beta e(C)})^{-1}} \tag{2.4}$$

L' equazione di stato si desume allora esprimendo $\lambda =$ in funzione della densità ρ e di β via:

$$\rho = \rho(\lambda,\beta) = \frac{1}{V}\sum_C \sum_n np(n;C) = \frac{1}{V}\sum_{n,C} \frac{ne^{-\beta\lambda n - \beta n e(C)}}{(1 - e^{-\beta\lambda - \beta e(C)})^{-1}} \tag{2.5}$$

e sostituendo poi λ con $\lambda(\beta,\rho)$ nella espressione della pressione (ricordando le proprietà dell' insieme gran canonico, si veda Insiemi Statistici,§5, (5.12)):

$$\beta p(\lambda,\beta) = \frac{1}{V}\log \Xi(\beta,\lambda) = -\frac{1}{V}\sum_C \log(1 - e^{-\beta\lambda - \beta e(C)}) \tag{2.6}$$

e della energia interna per unità di volume u_1:

$$\begin{aligned} u_1(\lambda,\beta) = \frac{1}{V}\sum_{C,n} n\, e(C)\, \frac{e^{-\beta(\lambda + e(C))n}}{(1 - e^{-\beta(\lambda+e(C))})^{-1}} = \\ = \frac{1}{V}\sum_C e(C) \frac{e^{-\beta(\lambda+e(C))}}{1 - e^{-\beta(\lambda+e(C))}} \end{aligned} \tag{2.7}$$

Per rendersi conto della differenza fra (2.5)/(2.7) e le proprietà del gas perfetto classico conviene immaginare che $e(C) = \underline{p}^2/2m$ se C è una celletta con centro nel punto $(\underline{p},\underline{q})$ e quindi di trascurare la variabilità di $\underline{p}^2/2m$ in C.

Questa approssimazione implica:

$$\begin{aligned} \beta p(\lambda,\beta) &= -\int \frac{\mathrm{d}^3\underline{p}}{h^3}\, \log(1 - e^{-\beta(\lambda - \underline{p}^2/2m)}) \\ \rho(\lambda,\beta) &= \int \frac{\mathrm{d}^3\underline{p}}{h^3}\, \frac{e^{-\beta(\lambda + \underline{p}^2/2m)}}{1 - e^{-\beta(\lambda+\underline{p}^2/2m)}} = 1/v \\ u_1(\lambda,\beta) &= \int \frac{\mathrm{d}^3\underline{p}}{h^3}\, \frac{\underline{p}^2}{2m}\, \frac{e^{-\beta(\lambda+\underline{p}^2/2m)}}{1 - e^{-\beta(\lambda+\underline{p}^2/2m)}} \end{aligned} \tag{2.8}$$

Integrando per parti la prima delle (2.8) si trova la relazione:

$$\beta p(\lambda,\beta) = 2\beta u_1/3 \tag{2.9}$$

L' aver trascurato la variabilità di $\underline{p}^2/2m$ in C introduce un errore; però non altera le proprietà qualitative delle (2.5)/(2.7) che ora metteremo in evidenza pur semplificando l' analisi grazie alla semplicità delle (2.8), (2.9).

Il fenomeno più rilevante è quello della "condensazione di Bose": le (2.4) mostrano che il parametro λ deve essere tale che $\lambda \geq \min_C e(C) = 0$. Dunque, come si vede dalla (2.8), la densità massima del sistema appare essere:

$$\rho_0(\beta) = \int \frac{\mathrm{d}^3\underline{p}}{h^3} \frac{e^{-\beta \underline{p}^2/2m}}{(1 - e^{-\beta \underline{p}^2/2m})} \tag{2.10}$$

che sembra errata perché la densità può essere prescritta a priori assegnando il numero di particelle, e non può essere limitata superiormente.

In realtà la densità può essere più elevata di $\rho_0(\beta)$ perché la (2.4) non ha senso se $\lambda \leq 0$, $e(C) = 0$. Intesa alla lettera la (2.4), per $\lambda > 0$, fa vedere che il numero di particelle in una delle cellette C con $e(C) = 0$ è:

$$\frac{\sum_{n=0}^{\infty} n e^{-\beta\lambda n}}{\sum_{n=0}^{\infty} e^{-\beta\lambda n}} = -\frac{\mathrm{d}}{\mathrm{d}\beta\lambda} \log \sum_{n=0}^{\infty} e^{-\beta\lambda n} =$$
$$= \frac{\mathrm{d}}{\mathrm{d}\beta\lambda} \log(1 - e^{-\beta\lambda}) = e^{-\beta\lambda}/(1 - e^{-\beta\lambda}) \xrightarrow[\lambda \to 0^+]{} \infty$$

L' interpretazione di (2.10) e della osservazione precedente è che le cellette con $e(C) > 0$ possono contribuire la quantità $\rho_0(\beta)$ alla densità ρ, al massimo: però il resto della densità, $\rho - \rho_0(\beta)$, è dovuto, se $\rho > \rho_0(\beta)$, alle particelle che si trovano nelle cellette C con $e(C) = 0$.

È questo il segno che il modo più appropriato di descrivere gli stati di densità $\rho > \rho_0(\beta)$ è di usare l' insieme canonico invece del gran canonico.

Però, dalla discussione ora fatta si può immaginare di descrivere uno stato di densità $\rho > \rho_0(\beta)$ nell' insieme gran canonico ponendo $\lambda = 0$ e immaginando che $(\rho - \rho_0(\beta))V$ particelle siano nelle cellette C con $e(C) = 0$.

Poiché $\rho_0(\beta) \to 0$ per $\beta \to \infty$ il fenomeno della condensazione di Bose è sempre importante a bassa temperatura, se la densità ρ vien mantenuta fissa. Ed è chiaro che le particelle che si trovano nella celletta C con $e(C) = 0$ hanno momento nullo e quindi non contribuiscono né all' energia interna n alla pressione né al calore specifico a volume costante.

In particolare volendo esaminare il calore specifico a volume costante per $T \to 0$ si può osservare che, non appena T è cos*piccolo*.ì. che $\rho_0(\beta) < \rho$ l' energia interna diviene:

$$U = u_1 V = V \int \frac{\mathrm{d}^3\underline{p}}{h^3} \frac{\underline{p}^2}{2m} \frac{e^{-\beta \underline{p}^2/2m}}{1 - e^{-\beta \underline{p}^2/2m}} = V\sigma T^{5/2} \tag{2.11}$$

ove:

$$\sigma = \int \frac{\mathrm{d}^3\underline{x}}{h^3} \underline{x}^2 \frac{e^{-\underline{x}^2}}{(1 - e^{-\underline{x}^2})^{-1}} (2m)^{5/3} k^{5/2} \tag{2.12}$$

ossia:

$$C_V = \frac{\partial u_1}{\partial T} = \text{ costante } T^{3/2} \quad \text{se } \rho > \rho_0(\beta) \tag{2.13}$$

che mostra come nel gas perfetto in esame non sia più vero che $C_v = 3nR/2$, bensì si ha $C_v \to 0$ per $T \to 0$! a bassa temperatura la equipartizione violata se si prende sul serio che $h \neq 0$.

Un altro esempio di situazione in cui h non può essere trascurato si ha quando si immagina che le particelle del gas interagiscano in modo molto semplice, anche se inconsueto nella Meccanica Classica, e precisamente in modo che non possano coesistere all' interno di una medesima celletta due o più particelle: la "stranezza" sta nel fatto che si deve immaginare che la forza dipenda dalle velocità oltre che dalla distanza, perché genera un cuore duro nello spazio delle fasi.

In questo caso la funzione di partizione è la (2.2) con la condizione che $n_C = 0, 1$:

$$\Xi(\lambda, \beta) = \prod_C (1 + e^{-\beta(\lambda + e(C))}) \tag{2.14}$$

e la probabilità che $n_C = n$ è, in luogo della (2.4):

$$p(n; C) = \frac{e^{-\beta(\lambda+e(C))n}}{(1 + e^{-\beta(\lambda+e(C))})} \qquad n = 0, 1 \tag{2.15}$$

e le (2.5)/(2.8) cambiano in conseguenza.

Questo gas è lungi dall' essere simile al gas perfetto classico e a bassa temperatura esibisce il fenomeno della "condensazione di Fermi": si vede invero che:

$$p(n; C) \xrightarrow[\beta\to\infty]{} \begin{cases} 1 & \text{se } e(C) < -\lambda \\ 0 & \text{se } e(C) > -\lambda \end{cases} \tag{2.16}$$

per cui a bassa temperatura sono occupate solo le cellette con $\underline{p}^2/2m < -\lambda$ i cui impulsi formano una sfera nello spazio degli impulsi ("sfera di Fermi"). Si noti che se λ 0 la densità del sistema tende a zero per $T \to 0$. Se si vuole mantenere costante la densità mentre $T \to 0$ occorre fissare $\lambda < 0$. Infatti se $\lambda < 0$ la densità è tale che:

$$\rho(\lambda, \beta) = \int \frac{\mathrm{d}^3\underline{p}}{h^3} \frac{e^{-\beta(\lambda+\underline{p}^2/2m)}}{(1 + e^{-\beta(\lambda+\underline{p}^2/2m)})} \xrightarrow[\beta\to\infty]{} \frac{4\pi}{3h^3}\sqrt{-2m\lambda}^3 \tag{2.17}$$

Dunque per $\beta \to \infty$ e densità costante, cioè $\lambda < 0$ costante, si trova l' energia interna ed il calore specifico a volume costante via le:

$$\begin{aligned} U &= V \int \frac{\underline{p}^2}{2m} \frac{\mathrm{d}^3\underline{p}}{h^3} \frac{e^{-\beta(\lambda+\underline{p}^2/2m)}}{(1 + e^{-\beta(\lambda+\underline{p}^2/2m)})} \\ C_V &\equiv \left(\frac{\partial U}{\partial T}\right)_V = V \int \frac{\underline{p}^2}{2m} \frac{\mathrm{d}^3\underline{p}}{h^3} k\beta^2 \frac{(\lambda + \underline{p}^2/2m)e^{-\beta(\lambda+\underline{p}^2/2m)}}{(1 + e^{-\beta(\lambda+\underline{p}^2/2m)^2})} \end{aligned} \tag{2.18}$$

ed uno studio degli integrali conduce, dopo alcuni calcoli, alla formula asintotica:

$$C_V \simeq \sigma V T \qquad \text{per } T \to 0 \tag{2.19}$$

con σ opportuna.

Dunque anche questo sistema si comporta in modo diverso dal gas perfetto classico a bassa temperatura: in particolare la (2.19) mostra che non si ha più equipartizione dell' energia (perché $C_V \neq V 3nR/2$)).

Le condizioni per distinguere il comportamento dei gas dell' uno e dell' altro tipo sopra considerati, in cui h non è più considerato trascurabile, rispetto al gas perfetto classico con il quale si confondono per $T \to \infty$, sono state discusse nella voce Insiemi Statistici, (2.2) e §6.

Ricordiamo che ivi si ottiene il valore della temperatura T al di sotto del quale gli effetti della grandezza delle cellette divengono importanti:

$$T_q = h^2/mk\rho^{-2/3} \ . \tag{2.20}$$

Si può verificare, sulle formule che precedono, che questo valore di T_q coincide, come ci si aspetta, col valore della temperatura tale che $\rho_0(\beta_q) \cong \rho$ nel primo caso e tale che $-\lambda\beta \cong 1$ nel secondo.

Si suol dire che la condizione $T > T_q$ è la condizione perché il gas perfetto non presenti fenomeni di "degenerazione" dovuti alla grandezza non trascurabile di h.

È facile ora rendersi conto che la degenerazione dovuta al fatto che h è positivo può essere il meccanismo che evita i paradossi dell' equipartizione dell' energia.

Ad esempio nella teoria di un cristallo il contributo degli elettroni al calore specifico è trascurabile perché per gli elettroni il valore della temperatura al di sotto del quale si manifestano i fenomeni di degenerazione (con conseguente piccolezza del calore specifico, (si veda (2.13) o (2.19)) può essere stimato sulla base della (2.20) e dà un valore molto grande per T_q.

Usando la (2.20) e $m = 0.91\ 10^{-27}$ g, $\rho = 10^{22}\text{cm}^{-3}$ (valore della densità degli elettroni liberi nel ferro) si trova T:

$$T_q \equiv 1/k\beta_q = 1.6\ 10^5\ ^\circ\text{K} \ . \tag{2.21}$$

Più in generale si può pensare che, se un dato sistema è fatto di varie particelle ciascuna con vari gradi di libertà interni, avvenga che, ad un certa data temperatura, solo alcuni gradi di libertà e alcune particelle siano non degeneri e fra loro si abbia equipartizione dell' energia; mentre altri gradi di libertà ed altre particelle siano degeneri con i conseguenti fenomeni, quali il loro "mancato" contributo al calore specifico.

Un esempio assai interessante è quello della teoria del corpo nero: invero il corpo nero è un sistema ad infiniti gradi di libertà, la maggior parte dei quali

si trovano in uno stato di *estrema degenerazione* per cui l' equipartizione dell' energia ha luogo solo fra un numero finito di gradi di libertà.

Nel §1 si è visto che una cavità riempita di radiazione elettromagnetica può essere pensata come un insieme di infiniti oscillatori armonici con hamiltoniana (1.14):

$$H = \frac{1}{2}\sum_{\alpha=1}^{2}\sum_{\underline{k}}(p^{(\alpha)}(\underline{k})^2 + c^2\underline{k}^2\underline{q}^{(\alpha)}(\underline{k})^2) \tag{2.22}$$

ove $p^{(\alpha)}(\underline{k})$ e $q^{(\alpha)}(\underline{k})$ sono variabili canoniche.

La distribuzione canonica attribuisce alle configurazioni in cui l' oscillatore di polarizzazione α e numero d'onda $\underline{k}$ si trova nella celletta $C^{\alpha,\underline{k}}_{m,n}$ con centro $(p^{(\alpha)}(\underline{k}), q^{(\alpha)}(\underline{k})) = (m\,\delta p, n\,\delta q)$, con m, n interi, la probabilità:

$$p(C^{\alpha,\underline{k}}_{m,n}) = \frac{e^{-\frac{\beta}{2}(m^2\delta p^2 + c^2\underline{k}^2\delta q^2 n^2)}}{\sum_{\overline{m},\overline{n}} e^{-\frac{\beta}{2}(\overline{m}^2\delta p^2 + c^2\underline{k}^2\overline{n}^2\delta q^2)}} \tag{2.23}$$

ove non si sono trascurate le dimensioni di $C^{\alpha,\underline{k}}_{mn}$ e si è preso sul serio l' insieme canonico (dimenticando che, come discusso alla fine del §2 della voce Insiemi Statistici, ove le dimensioni delle cellette non fossero trascurabili, in realtà la meccanica statistica stessa andrebbe completamente riformulata).

Ripercorrendo lo schema seguito per pervenire alla (2.20), si veda Insiemi Statistici, (2.2) e §6, si trova facilmente la condizione sotto la quale la dimensione di h è trascurabile; deve essere:

$$\sqrt{\beta}\delta p \ll 1, \qquad \sqrt{\beta}\; c|\underline{k}|\delta q \ll 1\ . \tag{2.24}$$

Senza fissare esplicitamente i valori di δp e δq si vede che le (2.24) implicheranno, in particolare, che β è troppo grande perché le dimensioni di h siano trascurabili se:

$$\beta c|\underline{k}|\delta p\,\delta q \equiv hc|\underline{k}| = \beta h\nu > 1 \tag{2.25}$$

ove $\nu = c|\underline{k}|/2\pi$, (ottenuta moltiplicando le (2.24)).

Ci si deve dunque attendere che, dato β, gli oscillatori con $|\underline{k}| \gg 1/hc\beta$ siano "degeneri", ossia *non possano essere descritti* senza tener conto delle dimensioni delle cellette.

Se $T = 6000\,°\mathrm{K}$ (temperatura della superficie del Sole) si trova che le frequenze degeneri sono tutte quelle di frequenza superiore a:

$$\nu_0 = 1/hc\beta = 1.25\ 10^{14}\ \text{cicli/sec} \tag{2.26}$$

se $h =$ costante di Plank; questo si può confrontare, a titolo di esempio, con la frequenza della luce verde (ove lo spettro solare ha il suo massimo) $\nu_{\text{verde}} = 0.6\ 10^{15}$ cicli/sec.

Questi valori numerici spiegano come il fenomeno della degenerazione sia stato particolarmente facile da osservare, e come mai tanto ruolo ha avuto nello sviluppo della meccanica quantistica.

L' energia media di un oscillatore non degenere è, per il principio di equipartizione dell' energia, $kT = 1/\beta$: mentre, non trascurando la possibilità di degenerazione, questa energia è:

$$u(\alpha, k) = \sum_{m,n} \frac{1}{2}(n^2\delta p^2 + c|\underline{k}|m^2\delta q^2)p(C_{mn}^{\alpha,\underline{k}})) \tag{2.27}$$

come espresso dalla (2.23).

La quantità di energia dovuta alla radiazione con frequenza fra ν e $\nu + d\nu$ è allora (si veda §1):

$$L^3 u_\nu \, \mathrm{d}\nu = (4\pi\nu^2/c^3)\, \mathrm{d}\nu L^3 \sum_{\alpha=1}^{2} u(k, \alpha) \tag{2.28}$$

ove $|\underline{k}| = 2\pi\nu/c$. Se $\nu \ll \frac{1}{\beta h}$ la (2.28) è semplicemente:

$$u_\nu = \frac{8\pi\nu^2}{c^3}\, \beta^{-1} \tag{2.29}$$

Per discutere il caso di alta frequenza $\nu \gg 1/hc$ occorre fissare δp e δq: ma in meccanica classica non si può dare un criterio chiaro per la scelta di δp o δq. Così, per concretezza, sceglieremo δp e δq in modo che:

$$\begin{cases} \delta p\, \delta q = h \\ \delta p = \vartheta^2 c|\underline{k}|\delta q \equiv \vartheta^2 2\pi\nu\delta q \end{cases} \Rightarrow \begin{cases} \delta p = \vartheta\sqrt{2\pi\nu h} \\ \delta q = \vartheta^{-1}\sqrt{h/2\pi\nu} \end{cases} \tag{2.30}$$

con $\vartheta \sim 1$ che, pur essendo una scelta naturale perché rende circa uguali i due addendi interni a (2.27) per $m = n = 1$ (esattamente uguali se $\vartheta = 1$), è pur sempre una scelta arbitraria. I risultati sono qualitativamente indipendenti dalla scelta di ϑ ma ne dipendono quantitativamente.

Dalle (2.28), (2.30), con una breve discussione della serie su m ed n, si vede che se $\beta h\nu \gg 1$:

$$\begin{aligned} u(\underline{k}, \alpha) &= \tilde{h}\nu e^{-\beta\tilde{h}\nu} \qquad \tilde{h} = \pi h\vartheta^2 \quad \text{se } \vartheta \neq 1 \\ u(\underline{k}, \alpha) &= 2\tilde{h}\nu e^{-\beta\tilde{h}\nu} \qquad \tilde{h} = \pi h \quad \text{se } \vartheta = 1 \end{aligned} \tag{2.31}$$

per cui la (2.29) fornisce (se $\vartheta \neq 1$ e $\tilde{h} = \pi h\vartheta^2$) la "legge di Wien":

$$u_\nu \cong \frac{8\pi\nu^2}{c^3}\, \tilde{h}\nu e^{-\beta\tilde{h}\nu} \tag{2.32}$$

che pur non potendo essere veramente presa sul serio (perché, come già osservato, per il fatto stesso che non trascuriamo le dimensioni di h, sarebbe necessario cambiare le leggi della meccanica) mostra che ad alta frequenza l' energia presente è di gran lunga inferiore al valore di equipartizione e, anzi, l' energia totale del campo elettromagnetico in equilibrio termico è finita, a differenza di quanto accadrebbe se ogni oscillatore avesse la stessa energia media.

Che la (2.32) non possa essere considerata corretta lo si vede anche dal fatto che cambiando poco la forma delle cellette, (si consideri (2.31) scegliendo $\vartheta = 1$ e $\vartheta \neq 1$), si troverebbe un risultato diverso.

Ad esempio Plank usò cellette a forma di corona ellittica definite da:

$$(n-1)h\nu \leq \frac{1}{2}\left(p^{(\alpha)}(\underline{k})^2 + c^2|\underline{k}|q^{(\alpha)}(\underline{k})^2\right) \leq nh\nu \qquad n \qquad \text{intero} \quad > 0 \tag{2.33}$$

e di area h: cioè immaginò che le cellette fossero definite dal valore della energia invece che dall' impulso e dalla posizione (questa è una forma che differisce "molto" da quella sopra considerata).

In questo modo (2.23) e (2.26) sono sostituite da :

$$\begin{aligned} p(C^{\alpha,\underline{k}}_{m,n}) &= \frac{e^{-\beta h\nu n}}{(1-e^{-\beta h\nu})^{-1}} \\ u(\underline{k},r) &= \sum_{n=0}^{\infty} nh\nu \frac{e^{-\beta nh\nu}}{(1-e^{-\beta h\nu})^{-1}} = \frac{h\nu e^{-\beta h\nu}}{1+e^{-\beta h\nu}} \end{aligned} \tag{2.34}$$

che conduce alla "formula di Plank":

$$\mu_r = \frac{8\pi\nu^2}{c^3}\frac{h\nu}{e^{\beta h\nu}-1} \tag{2.35}$$

per la radiazione del corpo nero.

Ovviamente sulla base della meccanica statistica classica è impossibile decidere quale sia la legge di radiazione corretta: si può solo dire che se effettivamente le cellette dello spazio delle fasi non possono essere scelte più piccole di una quantità minima, allora sarà impossibile accettare l' equipartizione, anzi, gli oscillatori con alta frequenza avranno energia media molto bassa.

Il fatto che la distribuzione sperimentale della radiazione in equilibrio termico (radiazione nera) sia conforme alla legge di Plank è un indice della realtà della non indefinita divisibilità dello spazio delle fasi. E il corpo nero è un sistema fisico molto adatto a mettere in evidenza la struttura discreta dello spazio delle fasi, perché consta di un numero infinito di oscillatori di frequenza ν che è più grande di un ν_0 comunque prefissato, (e quindi di un numero infinito di oscillatori degeneri se h è positivo, comunque piccolo).

Questo significa che la radiazione nera necessariamente contiene, a qualsiasi temperatura, oscillazioni per le quali le approssimazioni che conducono alla abituale trattazione statistica (in cui le cellette possono essere considerate puntiformi) non sono valide: e dunque lo studio teorico e sperimentale delle componenti di alta frequenza della radiazione nera permette di controllare la correttezza di varie ipotesi sulla meccanica dei punti in regioni dello spazio delle fasi ove le cellette di dimensione h non sono più correttamente approssimabili con punti.

§3 *meccanica statistica Quantistica. Introduzione.*

La meccanica statistica quantistica è assai simile, se vista da un punto di vista opportuno, a quella classica.

Lo spazio delle fasi non ha più senso e invece si pensa solo all' insieme delle grandezze osservabili: esse sono descritte da operatori lineari su uno spazio di Hilbert e gli insiemi statistici sono definiti in termini dell' operatore che descrive l' energia, di solito denotato con H e detto " operatore di Schrödinger".

Se A è un' osservabile e H descrive N particelle chiuse in un contenitore V, si definisce l' insieme canonico come l' insieme delle "matrici di densità" (che sono l' analogo delle misure di probabilità degli insiemi statistici classici) aventi la forma:

$$\rho = \ cost\, e^{-\beta H} \tag{3.1}$$

e il valore medio dell' osservabile A nello "stato" rappresentato da (3.1), che è parametrizzato da β e V come nel caso analogo dell' insieme canonico della meccanica statistica Classica, è definito da:

$$\overline{A} = \ \frac{Tr A\, e^{-\beta H}}{Tr\, e^{-\beta H}} \tag{3.2}$$

ove Tr è la traccia dell' operatore $\exp -\beta H$ (a tutti gli effetti pratici gli operatori possono essere pensati come "matrici di grandi dimensioni": dopo un po' di pratica si capisce infatti che tipo di attenzione occorre fare per evitare incresciosi errori).

La termodinamica si deduce ancora dalla funzione di partizione "canonica" quantistica:

$$Z(\beta, V) = Tr \exp -\beta H \tag{3.3}$$

e ora $1/\beta k$ è interpretato come la temperatura, mentre l' energia libera è definita da $f(\beta, v) = \lim_{V\to\infty} -\beta^{-1}\frac{1}{V}\log Z(\beta, V)$ nel limite $V \to \infty, V/N = v$ fisso (limite termodinamico).

Si noti che ora la temperatura assoluta *non è più definita* come proporzionale alla energia cinetica media, bensì ora *è identificata* come proporzionale al parametro $1/\beta$ che appare nella (3.2): si veda la voce Insiemi Statistici,§1, e il §1 di questa voce per un commento su questa differenza che in un certo senso è la novità più grande della meccanica statistica quantica.

Si possono anche definire insiemi statistici microcanonico e gran canonico e si può verificare formalmente, e talvolta rigorosamente, la loro equivalenza ai fini della termodinamica che essi definiscono.

Ad esempio, se si hanno N particelle identiche di massa m, racchiuse in un contenitore cubico V, lo spazio di Hilbert è $L^*_2(V^N)$ e l' operatore energia è:

$$H = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\Delta_{\underline{q}_i} + \Phi(\underline{q}) \tag{3.4}$$

ove $\Phi(\underline{q}) = \sum_{i<j}(\underline{q}_i - \underline{q}_j)$ il modello di energia potenziale di interazione, $\hbar = h/2\pi$, se h è la costante di Plank e $\Delta_{\underline{q}_i}$ è l' operatore di Laplace rispetto alle coordinate della i–ma particella (con opportune condizioni al contorno: ad esempio periodiche o "di Dirichlet"); l' asterisco in L^*_2 ricorda che lo spazio di Hilbert consta delle sole funzioni simmetriche o, in altri casi, antisimmetriche delle N coordinate di posizione $\underline{q}_1, \underline{q}_2, \cdots \underline{q}_N$.

La simmetria ("particelle di Bose–Einstein", " bosoni") o antisimmetria ("particelle di Fermi–Dirac"," fermioni") è imposta per tener conto della indistinguibilità delle particelle.

La nozione di stabilità dell' interazione Φ è importante nella meccanica statistica quantistica quanto in quella classica. Ora si formula al modo seguente: esiste $B \geq 0$ tale che l' operatore di Schrödinger H descrivente N particelle verifica, per ogni $N \geq 0$:

$$H \geq -BN \tag{3.5}$$

ove la disuguaglianza vale nel senso degli operatori.

È interessante osservare come ora la disuguaglianza (3.5) possa essere valida anche se $\inf \Phi(\underline{q}) = -\infty$ (in virtù del "principio di indeterminazione", si veda anche la voce "stabilità della materia").

Il caso di sistemi di particelle non identiche, come nella meccanica statistica classica, si riduce facilmente a quello delle particelle identiche.

Nel caso della meccanica statistica classica basta introdurre opportuni fattori combinatorii per tener conto della identità delle particelle, si veda Insiemi Statistici, §2, (2.19), (2.20); nel caso quantistico si richiederà la simmetria o antisimmetria delle funzioni d'onda, ossia degli elementi dello spazio di Hilbert, rispetto alle permutazioni delle posizioni delle *sole* particelle identiche.

Ad esempio un sistema di N_1 particelle cariche di carica $+e$ interagenti con la forza di Coulomb fra loro e con N_2 particelle di carica $-e$ avrebbe, nell' insieme canonico della meccanica statistica classica, la funzione di partizione:

$$Z(\beta, V) = \frac{1}{N_1! N_2} \int e^{-\beta\left(\sum_{i=1}^{N_1} \underline{p}_i^2/2m_+ + \sum_{i=1}^{N_2} \underline{p}_i^2/2m_-\right)} e^{-\beta\tilde{\Phi}(\underline{q})} \frac{\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}}{h^{3(N_1+N_2)}} \tag{3.7}$$

con

$$\begin{aligned}\tilde{\Phi}(\underline{q}) = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} \frac{-e^2}{|\underline{q}_i - \underline{q}_{j+N_1}|} + \frac{1}{2}\sum_{\substack{i,j=1\\ i\neq j}}^{N_1} \frac{e^2}{|\underline{q}_i - \underline{q}_j|} + \\ + \sum_{i,j=1}^{N_2} \frac{e^2}{|\underline{q}_{i+N_1} - \underline{q}_{j+N_1}|} + \sum_{i<j} \varphi(\underline{q}_i - \underline{q}_j)\end{aligned} \tag{3.8}$$

ove φ è un potenziale e descrive una forza non elettrica che si esercita fra le particelle, m_+ e m_- sono le masse delle particelle dei due tipi.

Nel caso quantistico si ha: $Z(\beta, V) = \mathrm{tr}\, \exp -\beta H$ ove H è l' operatore di Schrödinger:

$$H = -\frac{\hbar^2}{2m_+}\sum_{i=1}^{N_1} \Delta_{\underline{q}_i} - \frac{\hbar^2}{2m_-}\sum_{i=1}^{N_2} \Delta_{\underline{q}_{i+N_1}} + \tilde{\Phi}(\underline{q}) \tag{3.9}$$

considerato come operatore agente sullo spazio delle funzioni $f(\underline{q}_1, \ldots, \underline{q}_{N_1}, \underline{q}_{N_1+1}, \ldots, \underline{q}_{N_1+N_2})$ simmetriche o antisimmetriche rispetto alle permutazioni delle prime N_1 variabili o delle seconde N_2, ma senza alcuna proprietà di simmetria rispetto a permutazioni "miste".

La "statistica", ossia la simmetria delle funzioni dello spazio di Hilbert su cui è definita l' hamiltoniana gioca un ruolo essenziale nella teoria: già dal punto di vista classico del §1 si sono potuti intravedere i fenomeni che fanno sì che le statistiche quantistiche siano totalmente diverse anche da un punto di vista qualitativo, a basse temperature, (si veda il §2).

Ma la "statistica" gioca a volte un ruolo importante addirittura a temperature ordinarie.

Ad esempio i sistemi coulombiani senza altre forze che la forza di Coulomb sono *instabili* in meccanica statistica classica, a qualsiasi temperatura, per il motivo banale che il potenziale coulombiano fra particelle di carica opposta è inferiormente illimitato. Ma nella meccanica statistica quantistica sono stabili almeno se le particelle cariche verificano la statistica di Fermi–Dirac e se il sistema è elettricamente neutro (o anche se le particelle sono fermioni e bosoni e i bosoni hanno carica di uno solo dei due possibili segni): per una discussione di questo fatto e della sua osservabilità si veda la voce Stabilità della Materia .

§4 I problemi della meccanica statistica contemporanea

La meccanica statistica classica contemporanea si trova in una situazione ideale da un punto di vista concettuale: non sembrano esserci problemi teorici dopo che la compatibilità dei fenomeni irreversibili con la reversibilità della dinamica microscopica è stata capita nell' ottica della esistenza di varie scale di tempo e qualitativamente suffragata dai risultati rigorosi di Lanford sul limite di Grad–Boltzmann (si veda Meccanica Statistica Classica,§7), e dopo che i paradossi a cui la meccanica statistica classica conduce sono stati compresi in termini di effetti quantistici e le condizioni di applicabilità della meccanica statistica classica sono state corrispondentemente formulate in modo preciso (ad esempio si veda la (2.20) e le voci Meccanica Statistica Classica, (2.16), e Insiemi Statistici, (2.2) e §6).

Tuttavia, ancora dal punto di vista teorico, sarebbe necessario capire di più sulla dinamica dei sistemi di molte particelle fino, se possibile, a produrre algoritmi affidabili per il calcolo delle scale di tempo rilevanti nei fenomeni di avvicinamento all' equilibrio (si vedano le voci Coefficienti di Trasporto e Teorie Cinetiche).

Il problema ergodico è un problema a tutt'oggi poco compreso soprattutto in sistemi di particelle vicine a situazioni di equilibrio (come nelle oscillazioni dei cristalli) ove sembra che l' ipotesi ergodica sia per lo meno di dubbia validità come messo in luce dai moderni esperimenti numerici: si veda la voce "Equilibrio Dinamico–statistico". La scarsa comprensione quantitativa dei fenomeni di non equilibrio si traduce in grossi problemi nelle "Teorie Cinetiche" e delle teorie dei "Coefficienti di Trasporto" .

Ma i problemi aperti, e oggetto di intensa ricerca, non sono solo da ricercarsi nella meccanica statistica del non equilibrio.

Il problema centrale della meccanica statistica dell' equilibrio è forse quello della teoria delle transizioni di fase e dei relativi punti critici.

Non c'è alcuna evidenza dell' esistenza di problemi di fondo e si è recentemente chiarita la natura del fenomeno delle transizioni di fase, come fenomeno di " non equivalenza" degli insiemi statistici, ovvero di "sensibile dipendenza" dello stato di equilibrio dalle condizioni al contorno (si veda la voce Transizioni di Fase). E a mezzo di semplici modelli risolubili si è mostrato come anche i più semplici modelli di sistemi meccanici (come i sistemi magnetici di spin) possano presentare transizioni di fase non banali, e, anzi, assai interessanti. Tuttavia fenomeni così importanti come la transizione liquido–gas ovvero cristallo–liquido sono a tutt' oggi sostanzialmente non capiti.

Non esiste infatti alcun modello che possa essere trattato rigorosamente e che mostri queste transizioni, anche se la possibilità di tali transizioni e di una loro teoria fenomenologica, possono essere speculate ed elaborate sulla

base di innumerevoli teorie approssimate, la prima delle quali è la "teoria di campo medio", vedi voce, che fino agli anni '30 fu l' unica teoria disponibile per lo studio delle transizioni di fase. Quest' ultima è una teoria semplice ed interessante ma troppo approssimata (tanto da prevedere transizioni di fase liquido–gas in modelli unidimensionali per i quali si può rigorosamente provare che tali transizioni non possono aver luogo).

Recentemente la teoria delle transizioni di fase ha avuto sviluppi interessanti soprattutto per quel che riguarda la teoria dei fenomeni critici nel cui contesto sono state sviluppate nuove teorie, approssimate ma notevolmente più profonde della teoria del campo medio (si veda la voce "Gruppo di Rinormalizzazione").

Il fenomeno della "metastabilità", vedi voce, è un altro fenomeno assai mal compreso, anche perché sembra trattarsi di un fenomeno per la cui teoria è necessario combinare idee e metodi caratteristici sia dei problemi di evoluzione che dei problemi di equilibrio.

Un' altra classe di fenomeni mal compresi sono i fenomeni, di equilibrio e non, legati ai sistemi di particelle cariche, sistemi coulombiani, vedi voce: fino a pochi anni fa era persino poco chiaro come un sistema di particelle cariche ("materia") potesse stare in equilibrio termodinamico, data la grande intensità e lunga portata del potenziale coulombiano. Si avevano solo teorie fenomenologiche basate sulle stesse idee informanti la teoria di campo medio delle transizioni di fase già menzionata (in questo caso nota come teoria di Debye, vedi voce).

Recentemente il problema della stabilità della materia, vedi voce, stato risolto in modo soddisfacente nell' ambito della meccanica statistica quantistica (nella meccanica statistica classica si può dimostrare che il problema non ha soluzione, e "la materia è instabile" in assenza di forze altre che quelle coulombiane).

Resta però aperto il problema della comprensione quantitativa dell' equilibrio statistico di aggregati neutri di particelle cariche ("molecole") in un gas e dei fenomeni di "schermaggio", vedi voce, della forza coulombiana importanti nella fisica dei plasmi, vedi voce: ovviamente esistono teorie anche molto dettagliate di questi fenomeni, ma si vuol qui dire che non sono "fondamentali" e cioè che qua e la introducono ipotesi ulteriori giustificate solo sulla base euristico–fenomenologica e necessarie per superare ostacoli apparentemente insormontabili altrimenti, di natura matematica.

Ad esempio nella teoria dei gas molecolari di solito si postula che un dato sistema consista di particelle indistinguibili di caratteristiche date e invarianti nel tempo ("molecole") che interagiscono fra loro via forze effettive, risultato di interazioni elettromagnetiche schermate. È questa ovviamente una approssimazione in cui si ignora una parte importante del problema e cioè la possibilità che le molecole si dissocino, si ionizzino o reagiscano

chimicamente.

La meccanica statistica quantistica non solo risolve il problema concettuale della stabilità della materia, ma anche introduce la possibilità di comprendere teoricamente una grande varietà di nuovi fenomeni tipicamente legati alla natura quantica della Fisica microscopica: la "superfluidità" e la "superconduttività" ne sono i primi esempi, vedi voci.

Tuttavia si può dire che nella meccanica statistica quantistica sono presenti tutti i problemi teorici già presenti nella meccanica statistica classica, sono "più difficili" e sono in uno stato ancora più primitivo: ad eccezione, ovviamente, dei problemi riguardanti i fenomeni tipicamente quantistici, impossibili nella meccanica statistica classica.

La meccanica statistica infine è legata naturalmente a varie branche della matematica. La "teoria delle dtTransizioni di fase" e degli insiemi statistici ha dato nuova vita e nuovi problemi alla teoria dei processi stocastici e in generale alla teoria dei sistemi dinamici e dell' informazione. E ha fornito nuovi problemi (e a volte le relative soluzioni) allacombinatoria , ad esempio via i "modelli risolubili" e via la teoria dellapercolazione, vedi voci.

Hanno origine meccanico–statistica molti problemi di teoria delle equazioni differenziali alle derivate parziali e ordinarie e la meccanica statistica ha ispirato idee fondamentali nuove per lo studio di problemi apparentemente di natura diversa quali la turbolenza o la teoria dei campi quantizzati, (vedi voci).

In generale i problemi della meccanica statistica hanno dato luogo, a importanti problemi matematici al cui insieme, nella letteratura contemporanea, ci si riferisce con il nome di "formalismo termodinamico".

Si può dire che la MS si trova attualmente in un momento di particolare fioritura paragonabile alla situazione della meccanica al momento delle sue trionfali applicazioni nel '700 e nel primo '800. Nessun problema di fondo, nessuna contraddizione palese con esperimenti si è ancora manifestata e tuttavia restano ancora molti fenomeni semplici da interpretare e probabilmente interpretabili nell' ambito della teoria stessa: lo sforzo di comprensione di questi fenomeni ha spinto alla formulazione di problemi matematici nuovi e profondi e questo è un ulteriore segno di vitalità della MS: sembra infatti possibile affermare in modo non tautologico, che le teorie fisiche siano sempre sorgenti di problemi matematici importanti, ma solo finché sono realmente vive e con problemi non puramente tecnici da affrontare.

Modelli risolubili in Meccanica Statistica

Voce per il Dizionario delle Scienze Fisiche
dell' Istituto dell'Enciclopedia Italliana
Autore: Giovanni Gallavotti

Gennaio 1987

§1 *Introduzione.*

Dal 1921 sono stati scoperti vari modelli di sistemi della Meccanica Statistica che é stato possibile "risolvere esattamente". Il ruolo da loro svolto é stato essenziale, prima per dimostrare come la Meccanica Statistica classica potesse effettivamente prevedere l'esistenza di transizioni di fase e, successivamente, per fornire elementi di confronto e prova di metodi numerici sviluppati per lo studio delle transizioni di fase e della natura del punto critico.

In tutti i casi non banali si tratta di modelli unidimensionali o bidimensionali. Illustreremo qui i seguenti modelli:

1) Modello di Lenz–Ising a 1–dimensione.

2) Modello di Ising a due dimensioni, a campo magnetico nullo.
3) Modelli a vertice a 2 dimensioni.

In realtá l'ultimo modello, nella recente versione di Baxter, contiene i precedenti e gran parte degli altri modelli risolubili che qui non tratteremo; tuttavia é utile discutere qui, esplicitamente, anche i primi due.

Per quanto concerne la terminologia occorre precisare che quando si dice che un modello é risolubile esattamente di solito si intende che la sua energia libera o qualche altra funzione termodinamica é esplicitamente calcolabile, eventualmente in termini di una o piú quadrature (ossia in termini di un integrale finito–dimensionale, con dimensione indipendente dal numero di gradi di libertá del modello, di funzioni elementari). A volte si riescono a calcolare anche altre quantitá termodinamiche interessanti come alcune funzioni di correlazione o funzioni legate ad effetti di superficie. Tuttavia anche nei modelli risolubili rimangono, salvo in casi particolarmente banali, molte grandezze di interesse fisico, che tuttavia non si riesce a calcolare "esattamente".

§2 *Modello di Ising a* 1*–dimensione.*

Il modello di Ising 1–dimensionale ha valore puramente pedagogico e la sua teoria é (ormai) del tutto banale, e si trova in tutti i testi elemntari di Meccanica Statistica.

In generale il modello di Ising a d–dimensioni é definito in termini di un reticolo, per fissare le idee quadrato, Z^d costituito dai punti a coordinate intere nello spazio a d dimensioni. Se $\underline{\omega}$ é una regione parallelepipeda di Z^d, la si immaghina come il contenitore del sistema; e le configurazioni microscopiche si ottengono pensando che in ogni punto di $\underline{\omega}$ sia posto uno "spin" $\sigma = \pm 1$.

Dunque gli stati microscopici del sistema consistono di configurazioni $\underline{\sigma} \equiv (\sigma_i)_{i \in \underline{\omega}}$ e si suppone che la loro interazione sia descritta dalla funzione:

$$H(\underline{\sigma}) = -\Big(\sum_{<i,j>} J\sigma_i\sigma_j + \sum_i h\sigma_i \Big) \tag{2.1}$$

ove $\sum_{<i,j>}$ denota la somma sulle coppie di siti di $\underline{\omega}$ che sono primi vicini nel reticolo Z^d, e h e J sono parametri che rispettivamente hanno l'interpretazione di costante di accoppiamento e di campo magnetico esterno.

Il problema é quello di calcolare, data la funzione di partizione:

$$Z(\beta, J, \underline{\omega}) = \sum_{\underline{\sigma}} \exp -\beta H(\underline{\sigma}) \tag{2.2}$$

l' energia libera:

$$\lim_{\underline{\omega}\to\infty} -\frac{1}{|\underline{\omega}|}\log Z(\beta, J, \underline{\omega}) \equiv \beta f(\beta, J) \tag{2.3}$$

ove $\underline{\omega} \to \infty$ significa che il lato di $\underline{\omega}$ diviene grande restando $\underline{\omega}$ sempre di forma parallelepipeda, ad esempio, e centrato nell'origine.

La prima questione é se f sia o no analitica in funzione di β, h.

Nel caso $d = 1$ é facile svolgere il calcolo e si trova:

$$-\beta f_1(\beta, h) = \log\left(e^{\beta J}\cosh(\beta h) + (e^{-2\beta J} + e^{2\beta J}(\operatorname{senh}\beta h)^2)^{1/2}\right) \tag{2.4}$$

che é manifestamente una funzione analitica in h, β nella regione di interesse fisico (cioé per h, β reali); e quindi il modello 1–dimensionale di Ising non ha transizioni di fase.

Il ruolo di questo risultato é stato di far capire come le teorie di Van der Waals,di campo medio e diCurie–Weiss fossero criticabili come meccanismo per l' apparizione di transizioni di fase perché queste ultime teorie, applicate a questo modello, ne prevedevano l' esistenza *anche* ad 1 dimensione, nel caso "attrattivo" $J > 0$.

§3 *Il modello di Ising* 2*–dimensionale.*

Il caso $d = 2$ é molto piú interessante e la sua soluzione è fra i piú interessanti risultati della Fisica Teorica. E' risolubile solo in campo magnetico nullo ($h = 0$) e fu risolto da Onsager (1944) che mostró che (se $\beta J^* \equiv J^*(\beta)$ é definito da $\tanh J^*(\beta) = e^{-2\beta J}$ e se si denota $\cosh^{-1}$ la funzione inversa del cosh):

$$\begin{aligned}\beta f_2(\beta, 0) &= \frac{1}{2}\log 2\sinh(2\beta J) + \\ &+ \int_{-\pi}^{\pi}\frac{d\varphi}{4\pi}|\cosh^{-1}\left(\cosh 2J\cosh 2J^* + \sinh 2J\sinh 2J^*\cos\varphi\right)|\end{aligned} \tag{3.1}$$

e un semplice studio della dipendenza da β di questa funzione mostra che essa é singolare in corrispondenza del valore di $\beta = \beta_c$ per cui $J = J^*$ (ossia $\operatorname{senh}(2\beta J) = 1$) e la singolaritá si manifesta come una divergenza (logaritmica) della derivata di f_2 rispetto a β, e quindi del calore specifico.

Successivamente Onsager stesso(1948), e poi Yang–Yang, riuscirono anche a calcolare una stima per la "magnetizzazione" spontanea:

$$\lim_{h\to 0^+} -\frac{\partial\beta f_2}{\partial h}(\beta, h) = m(\beta) = \begin{cases} 0 & \text{se } \operatorname{senh}2\beta J < 1 \\ \left(1 - (\operatorname{senh}2\beta J)^{-4}\right)^{\frac{1}{8}} & \text{altrimenti}\end{cases} \tag{3.2}$$

congetturando che fosse il valore esatto, (come infatti fu dimostrato successivamente negli anni '70).

Questi sono i due risultati chiave che permisero di mostrare che la Meccanica Statistica poteva effettivamente prevedere transizioni di fase con esponenti critici diversi da quelli della teoria di campo medio: ad esempio dalla (3.2) si vede che $m(\beta) \xrightarrow[\beta\to\beta_c]{} 0$ come $(\beta-\beta_c)^{1/8}$ invece che come $(\beta-\beta_c)^{1/2}$ come previsto dalla teoria di campo medio).

Molte altre quantitá sono state calcolate (alcune giá nei lavori di Onsager e altre in lavori successivi); fra esse:

1) la funzione di correlazione $< \sigma_{\underline{O}}\sigma_{\underline{x}} >$ ove $\underline{O}$ denota l'origine delle coordinate e $\underline{x}$ é un punto del reticolo su uno dei due assi coordinati, ovvero si muove lungo la diagonale principale, e il simbolo $< \ldots >$ denota il valore medio della quantitá fra le $<>$, nella distribuzione di equilibrio di Gibbs.

Si dimostra che, se $\kappa(\beta) = 2\beta|J - J^*|$ e se $|\underline{x}|$ é grande, allora la funzione $< \sigma_{\underline{O}}\sigma_{\underline{x}} >$ é proporzionale a:

$$< \sigma_{\underline{O}}\sigma_{\underline{x}} > \sim \begin{cases} \dfrac{e^{-\kappa(\beta)|\underline{x}|}}{\sqrt{|\underline{x}|}} & \text{per } \beta < \beta_c \\ |\underline{x}|^{-1/4} & \text{per } \beta = \beta_c \\ \dfrac{e^{-\kappa(\beta)|\underline{x}|}}{|\underline{x}|^2} & \text{per } \beta > \beta_c \end{cases} \tag{3.4}$$

Molto recentemente é stato possibile determinare il comportamento asintotico di tutte le funzioni di correlazione di ordine arbitrario (ossia dei valori medi di prodotti dei valori degli spin in un numero arbitrario di siti) in vari regimi, ad esempio nel regime in cui $\beta \neq \beta_c$ e tutti i siti coinvolti nelle funzioni di correlazione si allontanano omoteticamente,Wu) (Cheng, McCoy, Wu).

2) La tensione superficiale fra fasi coesistenti, definita da:

$$\tau(\beta) = \lim_{\underline{\omega}\to\infty} \frac{1}{L} \log \frac{Z_{++}(\beta,\underline{\omega})}{Z_{+-}(\beta,\underline{\omega})} = \begin{cases} 0 & \text{se } \beta < \beta_c \\ 2\beta(J - J^*) & \text{se } \beta > \beta_c \end{cases} \tag{3.5}$$

ove Z_{++}, Z_{+-} denotano rispettivamente le funzioni di partizione del modello ottenuto fissando gli spin sul bordo di $\underline{\omega}$ uguali tutti a $+1$ nel primo caso, e uguali a $+1$ nella metá superiore e -1 in quella inferiore del bordo di $\underline{\omega}$, nel secondo caso; qui L é il perimetro di $\underline{\omega}$, supposto quadrato.

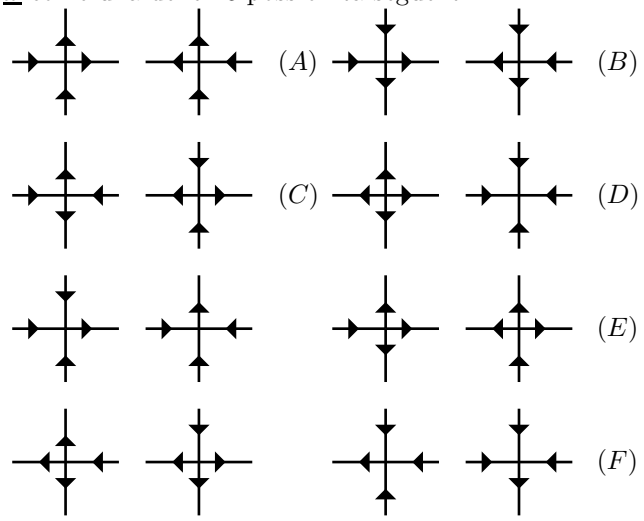
3) Molto é noto sulle proprietá dei valori medi di prodotti degli spin quando i loro siti vengono tenuti fissi sul bordo. Notevole é il fatto, ad esempio, che per $\beta \to \beta_c$ il valore medio di uno spin sul bordo (cioé la magnetizzazione spontanea sul bordo) non tenda a zero con l' esponente $1/8$ caratteristico della magnetizzazione in un punto interno, (a distanza fissa dall'origine, e

quindi infinitamente lontano dal bordo (nel limite in cui $\underline{\omega} \to \infty$)), ma con l'esponente 1/2.

§4 I modelli di vertice.

Si consideri una regione parallelepipeda $\underline{\omega} \subset Z^2$ con i lati opposti identificati (condizioni al bordo "periodiche"). Immaginiamo che gli stati microscopici del sistema si ottengano fissando un verso di percorrenza su ogni segmento ("legame") che collega primi vicini di $\underline{\omega}$.

Data una configurazione microscopica $\underline{\sigma}$ del sistema, questa apparirá in ogni sito di $\underline{\omega}$ come una delle 16 possibilitá seguenti:



Le configurazioni $\underline{\sigma}$ permesse nel "modello a otto vertici" sono le configurazioni, che denoteremo $S_8(\underline{\omega})$ tali che in ogni vertice le frecce appaiano come in A, B, C, D. Inoltre l'energia associata ad una configurazione é, nel modello generale degli otto vertici, una somma di contributi provenienti da ciascuno dei vertici j in modo che ciascuno contribuisca un'energia ε_j. I vertici A, B si chiamano per ragioni illustrate piú in basso "vertici polari" e quelli C, D "non polari".

Nel modello degli otto vertici risolubile (Baxter) si suppone che le energie dei vertici $1, 2$, cioè A, siano uguali fra loro e cosí pure quelle di $3, 4$, cioè B, di $5, 6$, cioè C, e di $7, 8$, cioè D, per cui il modello ha tre parametri indipendenti (ricordando che l'energia é definita a meno di una costante additiva se ne puó infatti eliminare uno).

E' un modello che puó essere equivalentemente formulato come un modello di spin che estende il modello di Ising, "generalizzato" con interazioni a due e quattro spin: si vede che posto $a = \exp(-\beta\varepsilon_A)$, $b = \exp(-\beta\varepsilon_B)$, $c =$

$\exp(-\beta\varepsilon_C)$, $d = \exp(-\beta\varepsilon_D)$) e definendo J, J', J'' in modo che:

$$\begin{aligned} a &= \exp\beta(J + J' + J'') \qquad b = \exp\beta(-J - J' + J'') \\ c &= \exp\beta(-J + J' - J'') \quad d = \exp\beta(J - J' - J'') \end{aligned} \tag{4.1}$$

il modello a otto vertici si identifica con quello di Ising con energia:

$$H(\underline{\sigma}) = \sum J\sigma_i\sigma_{i'} + \sum J'\sigma_i\sigma_{i''} + \sum J''\sigma_i\sigma_j\sigma_{j'}\sigma_{J''} \tag{4.2}$$

ove la somma é eseguita sui siti $i \in \underline{\omega}$ e i' denota il primo vicino di i sulla bisettrice del primo e terzo quadrante, i'' quello sulla bisettrice del secondo e quarto e j, j', j'' sono tre siti che con i formano un quadrato unitario con i in basso a sinistra.

Questo modello ha molti casi particolari interessanti, alcuni dei quali furono risolti prima del lavoro di Baxter (Lieb, Sutherland). Fra questi sono i modelli a 6 vertici, nelle cui configurazioni solamente sono permessi i vertici di tipo A, B, C e danno contributo $\varepsilon_A, \varepsilon_B, \varepsilon_C$ all'energia.

1) Il modello del "ghiaccio" di Pauling fissa $\varepsilon_A = \varepsilon_B = \varepsilon_C = 0$
2) Il KDP, (Slater), fissa le energie dei vertici B, C uguali ad $\varepsilon > 0$ e quelle dei vertici A nulle: i vertici "non polari" sono sfavoriti insieme a due, fra i quattro, vertici "polari".
3) Il modello F (Rys) fissa uguali a $\varepsilon > 0$ le energie dei vertici A, B e zero quelle di C: i vertici "polari" sono sfavoriti.

Tutti questi modelli hanno origine fisica legata alla teoria del " legame idrogeno" e dei legami chimici simili. Nel modello del ghiaccio i siti del reticolo rappresentano le posizioni di O nel cristallo e le direzioni delle frecce dicono dove si trovano i due H associati: se da un vertice emerge una freccia vuol dire che un atomo di H si trova in quella direzione piú vicino all' O dal quale emerge la freccia. Il fatto che le frecce siano associate ai legami e non ai siti consente appunto di imporre il vincolo che modella il legame idrogeno e che fa sì che il ghiaccio abbia una entropia residua inferiore a quella che avrebbe se nel cristallo di ghiaccio gli H potessero disporsi ciascuno accanto a "un" O (e così potrebbero generarsi configurazioni che hanno l'aspetto E, F o addirittura in cui su uno stesso legame si troverebbero frecce dirette in senso opposto).

Il modello KDP é apparso come modello delle proprietá ferroelettriche di KH_2PO_4 che cristallizza in tetraedri con al centro KPO_4 e i due H sulle linee che congiungono i KPO_4: un solo idrogeno puó trovarsi su ogni linea e puó trovarsi in due posizioni (e cioé vicino ad uno dei due estremi). Il KH_2PO_4 é un composto polare ma non ha simmetria sferica e pertanto non tutti i dipoli danno uguale contributo all'energia di una configurazione: i due vertici non polari C e due di quelli polari (ad esempio B) sono sfavoriti

e contribuiscono energia $\varepsilon > 0$ mentre gli altri contribuiscono $\varepsilon = 0$: a bassa temperatura ci si aspetta una polarizzazione spontanea: "ferroelettricitá".
Il modello F é invece un modello per una sostanza "antiferroelettrica" che almeno a basse temperature resisterá alla polarizzazione spontanea.
Si trova nel caso del modello del ghiaccio:

$$\lim_{\underline{\omega}\to\infty} \frac{1}{|\underline{\omega}|} \log Z(\underline{\omega}) = \frac{3}{2} \log \frac{4}{3} \tag{4.3}$$

Mentre nel caso del modello F l' energia libera é, se $\Delta = 1 - \frac{1}{2} e^{2\beta\varepsilon}$:

$$\begin{aligned} -\beta f_F(\beta) = &- \beta\varepsilon + \\ &+ \begin{cases} \frac{1}{8\mu} \int_{-\infty}^{\infty} \log \frac{\cosh\alpha - \cos 2\mu}{\cosh\alpha - 1} \frac{d\alpha}{\cosh \pi\alpha/\mu} & \text{se } \cos\mu = |\Delta| < 1 \\ \frac{\lambda}{2} + \sum_{n=1}^{\infty} \frac{e^{-n\lambda} \tanh n\lambda}{n} & \text{se } \cosh\lambda = -\Delta > 1 \end{cases} \end{aligned} \tag{4.4}$$

e nel caso del modello KDP, se $\Delta = \frac{1}{2} e^{\beta\varepsilon}$:

$$\begin{aligned} -\beta f_{KDP}(\beta) = \\ = \begin{cases} \frac{1}{8\mu} \int_{-\infty}^{\infty} \frac{\cosh\alpha - \cos\mu}{\cosh\alpha - \cos 3\mu} \frac{d\alpha}{\cosh \pi\alpha/2\mu} & \text{se } \Delta = -\cos\mu < 1 \\ 0 & \text{altrimenti} \end{cases} \end{aligned} \tag{4.5}$$

Inoltre i modelli F e KDP, a differenza dei modelli a otto vertici, sono risolubili anche in presenza di "campo elettrico" esterno E, se la presenza di tale campo é modellata supponendo che la energia di un vertice aumenti di $-Ep$ se p é il numero delle frecce verso l'alto meno quello delle frecce verso il basso. La soluzione in presenza di campo elettrico é in questi modelli molto semplice ma non la riportiamo per brevitá.
L'analisi delle espressioni ottenute conduce ai seguenti risultati che descriviamo denotando $f_F(\beta, E)$ e $f_{KDP}(\beta, E)$ le energie libere dei due modelli a temperatura β^{-1} e in presenza di campo elettrico E.
La funzione $f_F(\beta, 0)$ é infinitamente differenziabile ma ha una singolaritá essenziale, quale funzione analitica di β al valore $\beta = \beta_c$ in cui $\Delta = -1$: si ha dunque, a campo nullo, una transizione di fase di ordine ∞ . La polarizzazione, definita come numero medio di frecce verso l'alto è proporzionale alla derivata di f_F rispetto ad E, é nulla per $E = 0$ qualunque sia β; ma se $\beta > \beta_c$ (bassa temperatura) allora resta nulla anche se $E \neq 0$ e diviene non nulla solo se E passa un valore critico $E_c(\beta)$ e in questo senso il modello ha comportamento "antiferroelettrico". Se si tiene fisso E e si varia β si trova una transizione di fase di secondo ordine con singolaritá del calore specifico $\propto (\beta - \beta_c)^{-1/2}$.
Nel modello KDP la $f_{KDP}(\beta, E)$ é intrinsecamente diversa. Anche quì si ha una temperatura critica $\beta = \beta_c$ in campo E nullo (definita da $\Delta = 1$): in campo E nullo e se $\beta > \beta_c$ la polarizzazione vale 1 identicamente e l'energia

libera é costante; se $\beta \to \beta_c^-$ il calore specifico tende a zero come $(\beta_c - \beta)^{1/2}$ ma l'energia interna non tende a zero sebbene zero sia il valore dell'energia interna per $\beta > \beta_c$: dunque si ha una transizione di primo ordine con calore latente, e a bassa temperatura si ha polarizzazione spontanea ed il sistema é "congelato" ed ha funzioni termodinamiche banali. Inoltre per ogni campo elettrico E esiste una temperatura $T_c(E)$ tale che per temperature piú basse il sistema é polarizzato perfettamente ($p = 1$).

Queste proprietá, estratte dalle tante che si possono dedurre per questi modelli, ne mostrano la ricchezza della fenomenologia e l'interesse per la teoria delle transioni di fase, vista anche la differenza profonda che mostrano rispetto al modello di Ising.

Il modello a otto vertici, che sarebbe troppo lungo discutere quì, offre naturalmente una fenomenologia ancora piú interessante, oltre a presentare un enorme interesse teorico per la sofisticazione necessaria per pervenire alla sua soluzione. E' risolubile solo in campo elettrico nullo, ma anche in tal caso ha tre parametri indipendenti (come giá detto), ossia due oltre la temperatura.

Fissando questi parametri, arbitrariamente, il modello presenta sempre transizioni di fase: peró di tipo diverso secondo il valore di questi parametri.

L'interesse teorico, forse piú rilevante, é di permettere il calcolo di varie singolaritá al (esponenti critici)punto critico al variare delle energie dei vertici del modello, e di far vedere che tali esponenti sono funzioni continue non costanti di questi parametri. Anzi per opportuni valori dei parametri il modello si riduce al modello di Ising (o meglio a un sistema di due modelli di Ising indipendenti, c.f.r. le (4.1) con $J'' = 0$): per cui in un certo senso il modello di Baxter contiene "tutti" gli altri modelli risolubili.

In un momento in cui alcuni volevano credere all' ipotesi di universalitá intesa in senso troppo letterale (cioé che le transizioni di fase di "tutti i sistemi ragionevoli" producessero singolaritá critiche di uguale natura, ossia con uguali esponenti critici, o almeno che per esse ci fosse un numero piccolo di possibilitá) questa osservazione servì a mostrare l'esistenza di una infinitá continua di classi di universalitá. Ne risultó una assai migliore comprensione delle teorie del punto critico che proponevano spiegazioni teoriche dei fenomeni di universalitá, e prima fra esse quella del gruppo di rinormalizzazione, vedi voci corrispondenti.

Esistono molti altri modelli esattamente risolubili (modello "sferico" (Berlin, Kac), "dimeri" (Kastleyn), modello "XY", stato fondamentale del modello di Heisenberg a 1–dimensione (Yang, Baxter), di particolare interesse perché é alla sua soluzione che si puó ricondurre il problema dei sei e degli otto vertici, modelli di Ising su reticoli non quadrati, vari modelli 1–dimensionali *etc*): il lettore interessato potrá consultare le monografie sui modelli risolubili di Wu–McCoy sul modello di Ising e quella di Baxter

sui modelli risolubili più generali; una eccellente introduzione é ancora il lavoro di Lieb, Schultz, Mattis.

Riferimenti:

Mattis, Schultz, Lieb: Reviews of Modern Physics: 36, 856, 1964; B. McCoy, T.T.Wu, The two dimensional Ising model, Harvard University Press, Cambridge, 1973; R. Baxter, Exactly solved models, Academic Press, London, 1982.

æ

Limite termodinamico (problemi matematici).

voce per il DIZIONARIO DELLE SCIENZE FISICHE
edito dall' Istituto dell' Enciclopedia Italiana
Autore della voce Giovanni Gallavotti
Prima versione: agosto 1985
Versione 2: marzo 1986

§1 *Introduzione.*
§2 *Lo spazio delle fasi a volume infinito e le distribuzioni di probabilità su di esso.*
§3 *Caratterizzazione variazionale degli stati di Gibbs invarianti per traslazione.*
§4 *Caratterizzazione alternativa della distribuzione di Gibbs: equazioni DLR.*
§5 *Misure di Gibbs e processi stocastici.*

§1 *Introduzione*

La teoria degli insiemi statistici fornisce un modello di termodinamica solo nel limite di volume infinito, "limite termodinamico": è questa una situazione in cui si ha, anche, equivalenza fra i vari insiemi statistici (si vedano le voci Meccanica Statistica Clasica, §5, 6 e Insiemi Statistici, §3).

Gli elementi degli insiemi statistici descrivono le fasi termodinamiche del sistema in grande dettaglio, ben oltre la semplice definizione microscopica delle grandezze della termodinamica classica e la possibilità di calcolare le loro relazioni reciproche: ogni elemento di un insieme statistico descrive i dettagli delle configurazioni di equilibrio microscopiche tipiche della fase corrispondente.

Il problema del "limite termodinamico" è quello di stabilire un formalismo in cui rendere precise le affermazioni ora fatte e, quindi, di analizzarle a fondo in vari casi specifici. Qui considereremo solo il caso della meccanica statistica classica, in cui si può trascurare la grandezza della h di Planck, rinviando alle voci specifiche per le sue applicazioni. Come tutti i formalismi ha interesse solo perché nel suo ambito si possono agevolmente collocare applicazioni importanti che, però, potrebbero essere studiate in altri linguaggi con equivalente quantità di lavoro.

Tuttavia è ben noto che per ogni classe di problemi il formalismo in cui sono posti ha spesso un carattere unificatore e chiarificatore (si pensi al formalismo hamiltoniano in Meccanica) e la nascita di un formalismo "buono" di solito è successiva alla risoluzione di alcuni importanti problemi in un dato campo (ed è da essa suggerito). È così stato per il "formalismo del limite termodinamico" ed è in questa ottica che vanno lette le pagine che seguono.

Rendere precise le affermazioni sopra riportate nell' ambito di un formalismo chiaro consentirà poi di affrontare in modo coerente una gran quantità di problemi quali la descrizione delle correlazioni spaziali fra particelle in un gas, la descrizione di superfici di separazione fra fasi coesistenti, la formazione e dissociazione stazionaria di molecole ed altri fenomeni cooperativi.

La prima questione è in quale senso un elemento di un insieme statistico descriva una distribuzione di probabilità sullo spazio delle fasi una volta che si è considerato il limite di volume infinito. Considereremo qui solo l' insieme gran canonico per il quale la trattazione è un può più semplice. Si considera dunque un sistema di particelle identiche (per semplicità) di massa m racchiuse in un volume (cubico) V.

Supporremo che le particelle interagiscano via un potenziale φ che verifica almeno le condizioni di stabilità e temperatezza necessarie nella teoria degli insiemi statistici (si veda la voce Insiemi Statistici, §2, (2.18), (2.19)): ossia

$\Phi(\underline{q}_1, \ldots, \underline{q}_n) = \sum_{i<i} \varphi(\underline{q}_i - \underline{q}_j) \geq -Bn$ ("stabilità") e $|\varphi(\underline{r})| < C|\underline{r}|^{-(3+\varepsilon)}$, $B, C, \varepsilon > 0$, per $|\underline{r}| \geq r_0 > 0$.

Per evitare questioni di natura tecnica nella formulazione dei risultati supporremo anche che φ abbia un "cuor duro" di diametro r_0, cioè $\varphi(\underline{r}) = +\infty$ per $|\underline{r}| < r_0$: fisicamente questo significa che due particelle non possono avvicinarsi più di una distanza r_0. Ciononostante, con opportuni mutamenti eccezioni e problemi aperti, molto di quanto segue si applica anche al caso fisicamente forse più interessante e comune di "potenziali superstabili", ossia tali che esistano $B, A > 0$ per cui:

$$\Phi(\underline{q}_1, \ldots, \underline{q}_n) \geq -Bn + An^2/V \quad \text{se } \underline{q}_1, \ldots, \underline{q}_n \in V \tag{1.1}$$

se V è un volume cubico arbitrario contenente un numero arbitrario $n \geq 2$ di particelle in $\underline{q}_1, \ldots, \underline{q}_n$. Il potenziale di Lennard Jones è un tipico esempio di un potenziale superstabile, si veda la voce Viriale, (1.1), e Meccanica Statistica Classica, §2, (2.7).

Sia V un volume cubico e consideriamo l' elemento $\mu^{(\beta,\lambda,V)}$ dell' insieme gran canonico di parametri (β, λ) con particelle confinate in V : $\beta = 1/kT$, $k =$ costante di Boltzmann e $T =$ temperatura, $\lambda =$ "potenziale chimico", si veda la voce Insiemi Statistici, §5.

La probabilità di trovare n particelle nello stato microscopico $d\underline{p}_1 \ldots d\underline{p}_n$ $d\underline{q}_1 \ldots d\underline{q}_n$ secondo la distribuzione $\mu^{(\beta,\lambda,V)}$ è:

$$\begin{aligned} \mu^{(\beta,\lambda,V)}(d\underline{p}_1 \ldots d\underline{p}_n \, d\underline{q}_n \ldots d\underline{q}_n) = \\ = \frac{e^{-\beta(E(\underline{p},\underline{q})+\lambda n)}}{\Xi(\beta,\lambda)} \, \frac{d\underline{p}_1 \ldots d\underline{p}_n \, d\underline{q}_1 \ldots d\underline{q}_n}{n!} \end{aligned} \tag{1.2}$$

ove $E(\underline{p}, \underline{q}) = T(\underline{p}) + \Phi(\underline{q}) = \sum_{i=1}^n \underline{p}_i^2/2m + \Phi(\underline{q}_1, \ldots, \underline{q}_n)$ e Ξ è la funzione di partizione gran canonica (si veda la voce Insiemi Statistici, §5).

Vogliamo ora definire il limite di (1.2) per $V \longrightarrow \infty$ e interpretarlo come una distribuzione di probabilità sulle configurazioni del sistema infinito così ottenuto.

Il primo problema è di dare una definizione precisa delle configurazioni di un sistema infinito, poi di dare un significato preciso al limite di (1.2) per $V \longrightarrow \infty$, infine ci si porrà il problema della caratterizzazione delle distribuzioni che si ottengono in questo modo a partire dalla (1.2) o, più in generale, dalle analoghe distribuzioni gran canoniche con condizioni al contorno di particelle fisse (si veda Insiemi Statistici, §5).

§2 Lo spazio delle fasi a volume infinito e le distribuzioni di probabilità su di esso.

È naturale definire lo spazio delle fasi $\overline{M}$ a volume infinito come lo spazio delle successioni $(\underline{p},\underline{q}) = (\underline{p}_i,\underline{q}_i)_{i=1}^{\infty}$ di posizioni ed impulsi tali che in ogni volume finito ci sono solo un numero finito di particelle ("configurazioni localmente finite"): se si considerano sistemi di particelle con cuore duro di diametro $r_0 > 0$ basterà considerare le sole $\underline{q}$ tali che $|\underline{q}_i - \underline{q}_j| > r_0$, per $i \neq j$.

Però a causa della indistinguibilità delle particelle lo spazio delle configurazioni microscopiche sarà M, ottenuto da $\overline{M}$ identificando successioni $(\underline{p},\underline{q})$ che differiscono per una permutazione delle particelle.

Una "distribuzione di probabilità" μ su M è in generale definita in modo da permettere di dare una risposta alla domanda: quale è la probabilità che guardando in un dato volume V^0 (cubico per semplicità) si trovino in esso esattamente g particelle con impulsi in $d\underline{p}_1 \dots d\underline{p}_g$ e posizioni in $d\underline{q}_1 \dots d\underline{q}_g$?

Dunque la distribuzione μ sarà caratterizzata dalle funzioni $f_{V^0}(\underline{p}_1,\dots,\underline{p}_g, \underline{q}_1,\dots,\underline{q}_g)$ tali che la:

$$f_{V^0}(\underline{p}_1,\dots,\underline{p}_g,\underline{q}_1,\dots,\underline{q}_g)\; d\underline{p}_1 \dots d\underline{p}_g\; d\underline{q}_1 \dots d\underline{q}_g\,/g!$$

sia la probabilità ora descritta. Le f_{V^0} si diranno le "distribuzioni locali" di μ, ove è consuetudine non inglobare il fattore $g!$ nella f_{V^0} perchè, essendo le articele indistinguibili, questo consente varie semplificazioni dei calcoli combinatorii.

A mezzo delle funzioni f_{V^0} sarà possibile calcolare il valore medio di una "osservabile localizzata in V^0", ossia che dipende da $(\underline{p},\underline{q})$ solo attraverso lo stato delle particelle di $(\underline{p},\underline{q}) \in M$ che sono in V^0.

Se F è una tale osservabile si ha allora:

$$\begin{aligned}\overline{F} = \sum_{g=0}^{\infty} \int & F(\underline{p}_1,\dots,\underline{p}_g,\underline{q}_1,\dots,\underline{q}_g) \\ & f_{V^0}(\underline{p}_1,\dots,\underline{p}_g,\underline{q}_1,\dots,\underline{q}_g)\,\mathrm{d}p_1 \dots \mathrm{d}p_g\;\mathrm{d}q_1 \dots \mathrm{d}q_g/g!\end{aligned} \tag{2.1}$$

Data una distribuzione di probabilità, come la (1.2), descrivente un sistema di n particelle racchiuse in un volume "globale" V e fissato $V^0 \subset V$ (si pensi V come enorme e V^0 come molto più piccolo) si può calcolare la probabilità che all' interno di V^0 la configurazione $(\underline{p},\underline{q})$ appaia costituita di g particelle in $d\underline{p}_1 \dots d\underline{p}_g\, d\underline{q}_1 \dots d\underline{q}_g$: eseguiti gli opportuni integrali si troverà necessariamente un' espressione come:

$$f^{(V)}_{V^0}(\underline{p}_1,\dots,\underline{p}_g\,,\,\underline{q}_1,\dots,\underline{q}_g)\; d\underline{p}_1 \dots d\underline{p}_g\; d\underline{q}_1 \dots d\underline{q}_g/\,g!$$

È allora naturale definire il limite per $V \longrightarrow \infty$ (cioè quando si immagina di far tendere ad infinito il volume globale, ossia il contenitore del sistema)

della (1.2) come la distribuzione μ su M caratterizzata dalle distribuzioni locali:

$$f_{V^0}(\underline{p}_1, \dots, \underline{p}_g, \underline{q}_1, \dots, \underline{q}_g) = \lim_{V \to \infty} f^{(V)}_{V^0}(\underline{p}_1, \dots, \underline{p}_g, \underline{q}_1, \dots, \underline{q}_g) \qquad (2.2)$$

purché il limite esista per ogni V^0 fissato.

Si può dimostrare che se il potenziale interparticellare φ è superstabile, (1.1), e quindi anche se ha, allora il limite (2.2) esiste almeno su sottosuccessioni di $V \to \infty$.

E questo resta vero se $\mu^{(\beta,\lambda,V)}$ è sostituito con un più generale elemento di un insieme gran canonico con condizioni al contorno di particelle fisse fuori del contenitore globale V (si veda "Insiemi Statistici", §5, per questa nozione).

Definiamo allora l' insieme $G^0(\beta, \lambda)$ delle "distribuzioni di Gibbs" come l' insieme di tutte le possibili distribuzioni di probabilità che si ottengono come limiti di sottosuccessioni convergenti, nel senso (2.2), di distribuzioni gran canoniche $\mu^{(\beta,\lambda,V)}$ con condizioni al contorno diparticelle fisse.

Tali distribuzioni descrivono fasi di equilibrio termodinamico del sistema (si veda la voce Insiemi Statistici, §1, §5) e non sono necessariamente "invarianti per traslazione" cioè tali che per ogni spostamento $\xi \in R^3$:

$$f_{V^0+\xi}(\underline{p}_1, \dots, \underline{p}_g, \underline{q}_1 + \xi, \dots, \underline{q}_g + \xi) = f_{V^0}(\underline{p}_1, \dots, \underline{p}_g, \underline{q}_1, \dots, \underline{q}_g) \qquad (2.3)$$

infatti la invarianza per traslazioni è "rotta" dal fatto che il sistema è, prima del limite termodinamico $V \longrightarrow \infty$, contenuto in un cubo V; e non è detto che la invarianza venga restaurata per il solo fatto che $V \longrightarrow \infty$. Il fenomeno fisico dietro questa "rottura spontanea" (eventuale) della simmetria per traslazioni è la possibilità dell' esistenza di fasi termodinamiche in cui fasi pure diverse coesistono occupando, ad esempio, ciascuna metà dello spazio totale disponibile, separate da una superficie di separazione microscopicamente ben definita (si vedano le voci "Tensione Superficiale", "transizioni di fase", "rottura spontanea di simmetria"): si pensi ad un liquido in equilibrio con il suo vapore.

Distingueremo allora fra le distribuzioni di Gibbs $G^0(\beta, \lambda)$ il sottoinsieme $G(\beta, \lambda) \subset G^0(\beta, \lambda)$ delle distribuzioni di Gibbsinvarianti per traslazione, ossia con distribuzioni locali verificanti (2.3).

Se μ è una distribuzione di probabilità su M invariante per traslazioni e se $S = (S_1, S_2, S_3)$ sono le operazioni di traslazione delle configurazioni di M di una unità di lunghezza nelle direzioni x, y, z (ossia $S_\alpha(\underline{p}_i, \underline{q}_i)_{i=1}^{\infty} = (\underline{p}_i, \underline{q}_i + \underline{e}^\alpha)$ ove $\underline{e}^\alpha$, $\alpha = 1, 2, 3$, è il versore della direzione $\alpha-$ma), allora la terna (M, S, μ) è un sistema dinamico a tre dimensioni, si veda la voce Sistemi Dinamici.

Ovviamente si potrebbero definire le distribuzioni di Gibbs a partire da distribuzioni dell' insieme microcanonico (ovvero canonico) con condizioni al bordo di particelle fisse. E sorgerebbe il problema della equivalenza degli insiemi (si veda anche Insiemi Statistici, §3, 4, 5) nel senso che si dovrebbe verificare che la totalità delle distribuzioni di Gibbs costruite a partire dall' insieme gran canonico coincide con la totalità delle misure di Gibbs costruite a partire dagli insiemi canonico o microcanonico.

Questa verifica è difficile ed è pressoché completa solo nel caso di sistemi con cuore duro; ma è in larga misura incompleta nel caso "generale" di sistemi con potenziali superstabili. In ogni caso è un problema che qui non verrà discusso per evitare questioni troppo tecniche.

§3 Caratterizzazione variazionale delle distribuzioni di Gibbs invarianti per traslazioni

Consideriamo ora il problema di caratterizzare per altra via le misure di Gibbs $\mu \in G(\beta, \lambda)$, invarianti per traslazione, ("fasi pure omogennee").

La prima banale osservazione che segue direttamente dalla (2.2), o dalle sue varianti con diverse condizioni al contorno, è che se $\mu \in G^0(\beta, \lambda)$ allora la distribuzione degli impulsi è "maxwelliana", ossia $f_{V^0}(\underline{p}_1, \dots, \underline{p}_g, \underline{q}_1, \dots, \underline{q}_g)$ si può scrivere come:

$$f_{V^0}(\underline{p}_1, \dots, \underline{p}_g, \underline{q}_1, \dots, \underline{q}_g) = \frac{e^{-\beta \sum_{i=1}^{g} \underline{p}_i^2/2m}}{\sqrt{2\pi m \beta^{-1}}^{3g}} \overline{f}_{V^0}(\underline{q}_1, \dots, \underline{q}_g) \tag{3.1}$$

ove il fattore sotto radice è introdotto perché fornisce una normalizzazione evidentemente conveniente.

Le misure di probabilità sullo spazio delle fasi M le cui distribuzioni locali dipendono dagli impulsi come nella (3.1) si dicono distribuzioni "maxwelliane".

Il problema è, quindi, caratterizzare $\overline{f}_{V^0}$ in modo che la distribuzione definita dalle (3.1) sia in $G(\beta, \lambda)$.

Un ben noto argomento fa vedere che le (1.2) verificano un "principio variazionale". Precisamente se (1.2) viene scritta come:

$$f(\underline{p}_1, \dots, \underline{p}_n, \underline{q}_1, \dots, \underline{q}_n)\, \mathrm{d}\underline{p}_1 \dots \mathrm{d}\underline{p}_n\, \mathrm{d}\underline{q}_1 \dots \mathrm{d}\underline{q}_n \equiv f_n(\underline{p}, \underline{q})\, \mathrm{d}\underline{p}\, \mathrm{d}\underline{q}$$

e se:

$$E_n(\underline{p}, \underline{q}) \equiv \sum_{i=1}^{n} \underline{p}_i^2/2m - \Phi(\underline{q}_1, \dots, \underline{q}_n) \equiv T_n(\underline{p}) + \Phi_n(\underline{q}) \tag{3.2}$$

e se, ricordando che qui $(\underline{p},\underline{q})$ abbrevia $(\underline{p}_1,\dots,\underline{p}_n,\underline{q}_1,\dots,\underline{q}_n)$, si consideri il funzionale su f:

$$\begin{aligned} J(f) &= (1/V)\sum_{n=0}^{\infty}\int -f_n(\underline{p},\underline{q})\log f_n(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/n! - \\ &- (\beta/V)\sum_{n=0}^{\infty}\int f_n(\underline{p},\underline{q})(E_n(\underline{p},\underline{q})+\lambda n)\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/n! \qquad (3.3)\\ &= -(1/V)\sum_{n=0}^{\infty}\int f_n(\underline{p},\underline{q})(\log f_n(\underline{p},\underline{q})+\beta E_n(\underline{p},\underline{q})+\beta\lambda n)\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/n! \end{aligned}$$

Si vede subito che $J(f)$ è stazionario (massimo) nell' insieme delle $f \geq 0$ tali che:

$$\sum_{n=0}^{\infty} f_n(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/n! = 1$$

se, con la notazione $(\underline{p},\underline{q}) = (\underline{p}_1,\dots,\underline{p}_n,\underline{q}_1,\dots,\underline{q}_n)$, f_n verifica per ogni $n,\underline{p},\underline{q}$:

$$(-\log f_n(\underline{p},\underline{q}) - \beta(E_n(\underline{p},\underline{q})+\lambda n)) = \text{ costante} \qquad (3.4)$$

e, cioè se f è data dalla (1.2).

Viene allora spontaneo definire sullo spazio di tutte le distribuzioni invarianti per traslazione μ su M e che hanno la distribuzione degli impulsi " maxwelliana" i seguenti funzionali: volume specifico delle particelle, energia potenziale corrispondente al potenziale interparticellare φ e entropia.

Questi funzionali, che denoteremo rispettivamente $v(\mu), u(\mu)$ ed $s(\mu)$, li scriveremo, prima nel caso generale e poi nella espressione che assumono se f_{V^0} ha la forma (3.1).

Abbreviamo la notazione anche per le funzioni di distribuzione locali nel volume V^0 come:

$$\begin{aligned} f_{V^0,g}(\underline{p},\underline{q}) &\equiv f_{V^0}(\underline{p}_1,\dots,\underline{p}_g,\underline{q}_1,\dots,\underline{q}_g)\\ \overline{f}_{V^0 g}(\underline{q}) &= \overline{f}_{V^0}(\underline{q}_1,\dots,\underline{q}_g) \end{aligned}$$

ove $(\underline{p},\underline{q})$ sta per $(\underline{p}_1,\dots,\underline{p}_g,\underline{q}_1,\dots,\underline{q}_g)$ e $\mathrm{d}\underline{p}\,\mathrm{d}\underline{q} = \mathrm{d}\underline{p}_1\dots\mathrm{d}\underline{p}_g\,\mathrm{d}\underline{q}_1\dots\mathrm{d}\underline{q}_g$. Allora il volume specifico della distribuzione maxwelliana μ sarà definito come:

$$\begin{aligned} v(\mu)^{-1} &= \lim_{V^0\to\infty}(1/V^0)\sum_{g=0}^{\infty}\int g f_{V^0,g}(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/g! = \\ &= \lim_{V^0\to\infty}(1/V^0)\sum_{g=0}^{\infty}\int g\overline{f}_{V^0,g}(\underline{q})\,\mathrm{d}\underline{q}/g! \qquad (3.5) \end{aligned}$$

L' energia totale di energia potenziale φ sarà, usando la notazione (3.2):

$$\begin{aligned} u_\varphi(\mu) &= \lim_{V^0\to\infty}(1/V^0)\sum_{g=0}^{\infty}\int (T_g(\underline{p})+\Phi_g(\underline{q}))f_{V^0,g}(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/g! \\ &= \lim_{V^0\to\infty}(1/V^0)\sum_{g=0}^{\infty}\int (3g/2\beta+\Phi_g(\underline{q}))\overline{f}_{V^0,g}(\underline{q})\,\mathrm{d}\underline{q}/g! \qquad (3.6)\\ &\equiv (3/2\beta)v(\mu)^{-1}+\overline{u}_\varphi(\mu) \end{aligned}$$

ove nel penultimo passaggio si sono eseguiti esplicitamente gli integrali, gaussiani e quindi banali, sulle $\underline{p}$ e nell' ultimo si è usata la (3.5).

Similmente l' entropia termodinamica è:

$$\begin{aligned} s(\mu) &= \lim_{V^0\to\infty} -(1/V^0)\sum_{g=0}^{\infty}\int f_{V^0,g}(\underline{p},\underline{q})\log f_{V^0,g}(\underline{p},\underline{q})\,\mathrm{d}\underline{p}\,\mathrm{d}\underline{q}/g! \\ &= \lim_{V^0\to\infty} -(1/V^0)\sum_{g=0}^{\infty}\int \frac{e^{-\beta T_g(\underline{p})}}{\sqrt{2\pi m\beta^{-1}}^{3g}}\,\overline{f}_{V^0,g}(\underline{q}) \\ (-\beta T_g(\underline{p})-(3g/2)\log 2\pi m\beta^{-1}+\log\overline{f}_{V^0,g}(\underline{q}))\,\mathrm{d}\underline{q}/g! &= \qquad (3.7)\\ &= \lim_{V^0\to\infty} -(1/V^0)\sum_{g=0}^{\infty}\int f_{V^0,g}(\underline{q})\,\mathrm{d}\underline{q}/g! \\ (-3g/2\beta-(3g/2)\log\,2\pi m\beta^{-1}+\log\overline{f}_{V^0,g}(\underline{q}))&\equiv \\ &\equiv -(3/2)v(\mu)^{-1}-v(\mu)^{-1}\,\log\,(2\pi\,e\,m\beta^{-1})^{3/2}+\overline{s}(\mu)\ . \end{aligned}$$

Si può dimostrare che nel caso di sistemi a cuore duro, tutti i limiti sopra descritti esistono: esistono però assai più in generale ma non discuteremo oltre questo punto.

Si consideri ora il problema di massimizzare sullo spazio delle distribuzioni μ su M, invarianti per traslazione e con distribuzioni locali maxwelliane (cioè aventi la forma (3.1)), il funzionale:

$$s(\mu)-\beta\lambda v(\mu)^{-1}-\beta u_\varphi(\mu) \qquad (3.8)$$

Detto $\beta P(\beta,\lambda)$ l' estremo superiore di (3.8) citeremo solo i risultati validi nel caso di particelle con cuore duro per evitare discussioni tecniche sul caso superstabile più generale (1.1) (caso in cui i risultati non sono ancora completamente soddisfacenti da un punto di vista matematico).

Nel caso dunque che φ abbia cuore duro si ha:

$$\beta P(\beta,\lambda)=\max_{\mu}(s(\mu)-\beta\lambda v(\mu)^{-1}-\beta u_\varphi(\mu)) \qquad (3.9)$$

e il massimo è raggiunto esattamente su tutte e sole le distribuzioni di Gibbs $\mu \in G(\beta, \lambda)$ invarianti per traslazione. Si può mostrare che il massimo P ha il significato termodinamico di "pressione", si veda Insiemi Statistici, §2.

Quello che di notevole deve essere sottolineato nella proprietà variazionale (3.9) è che, sebbene sia stata euristicamente fondata sul funzionale (3.8), che conduce all' elemento (β, λ) dell' insieme gran canonico a volume V del contenitore, finito e senza condizioni al contorno, accade tuttavia che le soluzioni del problema di massimo (3.8) contengono anche le distribuzioni di Gibbs che si ottengono imponendo a volume finito condizioni al bordo generali di particelle fisse e poi considerando il limite termodinamico.

Si può inoltre dimostrare che l' insieme delle misure μ che realizzano il massimo in (3.9) (e cioè, per quanto detto, $G(\beta, \lambda)$) sono un insieme convesso (ossia, $\mu_1, \mu_2 \in G(\beta, \lambda)$ implica $a\mu_1 + (1-a)\mu_2 \in G(\beta, \lambda)$ per ogni $a \in (0, 1)$).

E, di più, questo insieme convesso è in realtà un simplesso, ossia tale che ogni $\mu \in G(\beta, \lambda)$ può essere rappresentato in un unico modo come sovrapposizione convessa di punti estremali di $G(\beta, \lambda)$.

Il significato dell' ultima poprietà descritta è interessante: dice che se le distribuzioni estremali di $G(\beta, \lambda)$ sono interpretate come le fasi pure omogenee (cioè invarianti per traslazione) allora tutti gli altri elementi di $G(\beta, \lambda)$ sono loro miscugli e possono essere rappresentati come tali in un solo modo.

Ad esempio se $G(\beta, \lambda)$ ha solo due punti estremali μ_+ e μ_- (l' uno rappresentante la fase "liquida" e l' altro la fase "gas"), ogni altra distribuzione in $G(\beta, \lambda)$ può essere rappresentata come $a\mu_+ + (1-a)\mu_-$ con $0 < a < 1$, ove a ha l' interpretazione di frazione di massa nella fase liquida.

È infine notevole che si possa dimostrare che gli stati estremali μ di $G(\beta, \lambda)$ godono della proprietà di "ergodicità" nel senso che i sistemi dinamici (M, S, μ) sono "ergodici", (si vedano le voci Sistemi Dinamici, Entropia ed Informazione), e sono i soli elementi di $G(\beta, \lambda)$ con questa proprietà. Il significato fisico è che in una fase pura omogenea le particelle che si trovano in due regioni cubiche molto distanti tra loro sono distribuite indipendentemente almeno in media (cioè se si media la loro distribuzione sulla distanza fra i centri delle due regioni). Questa proprietà che intuitivamente dovrebbe caratterizzare le fasi pure è quella su cui si fonda l' interpretazione delle distribuzioni estremali di $G(\beta, \lambda)$ come fasi pure omogenee.

§4 Caratterizzazione alternativa delle distribuzioni di Gibbs: equazioni DLR.

A mezzo del principio variazionale (3.9) si trovano tutte le distribuzioni di

Gibbs invarianti per traslazione. Su basi fisiche tuttavia ci si aspetta, come già detto, che possano esistere anche distribuzioni di Gibbs non invarianti per traslazione: ossia, nelle notazioni del §3, in generale si avrà che $G^0(\beta,\lambda)$ contiene $G(\beta,\lambda)$, ma non coincide con $G(\beta,\lambda)$.

Allora conviene ricercare anche un'altra caratterizzazione degli stati di Gibbs che "non discrimini" gli stati non invarianti per traslazione.

Una tale caratterizzazione è possibile e nasce ancora da un ragionamento euristico basato sulla formula (1.2) della distribuzione $\mu^{(\beta,\lambda,V)}$ gran canonica a volume finito V e senza particelle fisse fuori di V.

Ci si domanda: dato $V^0 \subset V$ quale è la probabilità di trovare in V^0 esattamente g particelle nelle posizioni $\underline{q}_1,\dots,\underline{q}_g$ sapendo che fuori di V^0 le particelle sono nelle posizioni $\underline{q}'_1,\underline{q}'_2,\dots$?

Denotata $\overline{f}_{V^0}(\underline{q}_1,\dots,\underline{q}_g/\underline{q}'_1,\underline{q}'_2,\dots) = \overline{f}_{V^0 g}(\underline{q}/\underline{q}')$, se qui $\underline{q}$ abbrevia $(\underline{q}_1,\dots,\underline{q}_g)$ e $\underline{q}'$ abbrevia $(\underline{q}'_1,\underline{q}'_2,\dots)$, la densità di questa "probabilità condizionata", è immediato dedurre da (1.2) che:

$$\begin{aligned}\overline{f}_{V^0}(\underline{q}_1,\dots,\underline{q}_g/\underline{q}'_1,\underline{q}'_2,\dots) &\equiv \overline{f}_{V^0,g}(\underline{q}/\underline{q}') = \\ &= \frac{\exp\left[-\beta\lambda g - \beta\Phi(\underline{q}_1,\dots\underline{q}_g) - \beta\sum_{i=1}^{n}\sum_{j\geq 1}\varphi(\underline{q}_i-\underline{q}'_j)\right]}{\text{normalizzazione}}\end{aligned} \tag{4.1}$$

e la normalizzazione si determina imponendo che $\overline{f}_{V_0}$ definisca una distribuzione di probabilità, cioè che:

$$\sum_{g=0}^{\infty}\int \overline{f}_{V^0 g}(\underline{q}/\underline{q}')\,\mathrm{d}\underline{q}/g! = 1 \tag{4.2}$$

Questa relazione dipende dal volume V totale solo perché $\underline{q}'_1,\underline{q}'_2\dots$, ossia le particelle della configurazione esterna a V^0, devono essere contenute in V.

Viene quindi naturale definire, in alternativa ai §2,3, una distribuzione di Gibbs su M di parametri (β,λ) come una distribuzione μ su M maxwelliana sugli impulsi e per la quale la probabilità condizionata per l' evento in cui le posizioni delle particelle in V^0 sono $\underline{q}_1,\dots,\underline{q}_g$ sapendo che fuori di V^0 le particelle sono in $\underline{q}'_1,\underline{q}'_2\dots$ è data dalla (4.1) senza più la condizione che $\underline{q}'_1,\underline{q}'_2,\dots$ siano nel contenitore globale V (che ora si immagina infinito).

Così letta la (4.1) va sotto il nome di "equazione DLR" (Dobrushin, Lanford, Ruelle).

Il teorema naturale è che le distribuzioni di Gibbs, definite come soluzioni della (4.1) con distribuzione maxwellinana sugli impulsi, sono esattamente l' insieme $G^0(\beta,\lambda)$ delle distribuzioni di Gibbs, definite via il limite termodinamico nei paragrafi precedenti e che siano o no invarianti per traslazione.

La dimostrazione di questo teorema è completa solo nel caso di sistema di particelle con cuore duro, ma ne esistono vari adattamenti ad altri casi, incluso il caso (1.1) dei potenziali superstabili.

§5 Misure di Gibbs e processi stocastici.

Le distribuzioni μ di probabilità sullo spazio M definiscono, per integrazione sulle coordinate $\underline{p}$ di impulso, corrispondenti distribuzioni di probabilità sullo spazio M_1 delle sole configurazioni delle posizioni $\underline{q}$ di infinite particelle. E le misure di Gibbs integrate sulle coordinate di impulso (a banale distribuzione maxwelliana) forniscono una interessante classe di misure su M_1, che chiameremo ancora distribuzioni di Gibbs.

In generale le distribuzioni μ di probabilità su M_1 sono note in teoria della probabilità come "processi stocastici puntuali" perché il punto $\underline{q} \in M_1$ descrive in realtà una famiglia di particelle localizzate nei punti $\underline{q}_1, \underline{q}_2, \ldots$ in R^3 se $\underline{q} = (\underline{q}_1, \underline{q}_2, \ldots)$.

Questa osservazione permette di dare una nuova interpretazione fisica a vari risultati della teoria dei processi stocastici puntuali e soprattutto consente di tradurre problemi importanti per la Fisica in interessati problemi matematici della teoria dei processi puntuali.

Quella che forse è la questione centrale è di far vedere che esistono scelte semplici del potenziale interparticellare φ, supposto anche a cuore duro, in corrispondenza delle quali il principio variazionale o le equazioni DLR ammettono più di una soluzione per opportuni valori di λ e se β è abbastanza grande (cioè se la temperatura è abbastanza bassa).

Questo è il problema della esistenza della transizione liquido–gas in un sistema omogeneo di particelle identiche: per il quale ancora non esiste un esempio trattato in modo matematicamente rigoroso (ossia in cui non vengano introdotte ulteriori ipotesi *ad hoc* al "momento opportuno").

Altri problemi notevoli sono quelli connessi a questioni di invarianza di scala.

Dalla Fisica Teorica viene l' ipotesi che non solo la transizione liquido gas abbia luogo non appena φ oltre al cuore duro abbia una "coda" attrattiva (cioè negativa), ma anche che tale transizione abbia un punto critico (λ_c, β_c) in cui la distribuzione di Gibbs (ed il processo stocastico corrispondente) μ ha particolari proprietà di scala.

Precisamente si immagini di pavimentare R^3 con cubi $Q_{\underline{n}}$ di lato L parametrizzati da tre interi $\underline{n} = (n_1, n_2, n_3)$ in modo che il cubo $Q_{\underline{n}}$ sia fatto dei punti di R^3 con coordinate $n_h L \le x_h < (n_h + 1)L$, $h = 1, 2, 3$.

Si definisca la famiglia di variabili (cioè di funzioni) $\sigma_{\underline{n}}$ su M:

$$\sigma_{\underline{n}} = [(\text{numero di particelle in } \ Q_{\underline{n}}) - v(\mu)^{-1} L^3]/L^{3\delta/2} \qquad (5.1)$$

ove δ è un parametro da scegliere.

Si ottiene un processo stocastico i cui "stati sul sito $\underline{n}$" sono variabili reali indicizzate da $\underline{n}$ e definite dalle (5.1).

Si ritiene, sulla base della teoria del "punto critico", vdei voce, che, nel limite $L \to \infty$ e se δ è scelto opportunamente, il processo stocastico che descrive la distribuzione delle variabili $\sigma_{\underline{n}}$ tenda ad un processo stocastico "omogeneo" nel senso che la distribuzione limite che descrive le variabili $\sigma_{\underline{n}}$ è tale che esse possano essere rappresentate come:

$$\sigma_{\underline{n}} = \int_{Q_{\underline{n}}} \psi(x)\,\mathrm{d}x \tag{5.2}$$

ove $\psi(x)$ è un processo stocastico su R^3 (a valori reali) con "funzioni di correlazione" omogenee. Cioè per ogni k ed ogni $x_1, x_2, \ldots, x_k$:

$$\mathcal{E}(\psi(x_1)\ldots\psi(x_k)) = \text{ funzione omogenea di } (x_1, \ldots, x_k) \tag{5.3}$$

se $\mathcal{E}$ denota l' operazione di calcolo del valore medio ("aspettazione" nel linguaggio dei processi stocastici) rispetto alla distribuzione delle variabili aleatorie ψ.

Poiché non sono noti esempi "non banali" di processi stocastici di questo tipo si comprende qui l' interesse anche matematico della teoria delle transizioni di fase, che nei suoi aspetti euristici fornisce soluzioni a vari problemi connessi (fra i quali appunto quello della esistenza e costruzione di processi stocastici omogenei non "banali", si veda gruppo di rinormalizzazione , punti fissi non banali, si vedanole voci), che si tramutano in importanti congetture matematiche ed idee per la loro dimostrazione.

Ma questi non sono che primi esempi delle ragioni dell' enorme interesse suscitato nella teoria delle probabilità dalla teoria delle distribuzioni di Gibbs e del limite termodinamico: si può dire che la teoria dei processi stocastici è stata, negli ultimi decenni, letteralmente rivoluzionata dagli influssi della Meccanica Statistica.

Concludiamo menzionando che i metodi e i risultati della teoria degli stati di Gibbs hanno recentemente avuto importanti applicazioni alla teoria della turbolenza e dei moti caotici: si vedano le voci Turbolenza, Sistemi Dinamici, Entropia ed Informazione, Simulazione Numerica nei Fluidi, Caos.

æ

Moto Browniano.

per il "DIZIONARIO DELLE SCIENZE FISICHE"
edito dall' Istituto della Enciclopedia Italiana
Autore della voce: Giovanni Gallavotti
31 marzo 1991

§1 *Il moto browniano e la teoria di Einstein.*
§2 *La teoria di Smoluchowski.*
§3 *La teoria di Uhlenbeck Ornstein.*
§4 *La teoria di Wiener.*

§1 Il moto browniano e la teoria di Einstein.

Fu osservato per la prima volta da R. Brown, nel 1828, il quale riconobbe che il moto di "molecole" di un certo polline ("colloide") di grandezza di circa 10^{-3} cm in soluzione non era dovuto a cause interne o biologiche, come fino ad allora si credeva, ma aveva natura meccanica (sconosciuta) ed era comune a particelle di simile taglia di qualunque materiale, organico o inorganico. Ben presto tali moti vennero attribuiti a urti con i costituenti microscopici della materia (tra i primi a riconoscere questo fu G. Cantoni (1868)). Il moto browniano destò l' interesse di molti, fra cui H. Poincaré; e la teoria del moto browniano è dovuta ad A. Einstein e, indipendentemente, a M. Smoluchowski, (1905-1906), presto seguita dalle conferme sperimentali di J. Perrin, (1908).

La principale critica (K. Nägeli, 1879) all' ipotesi sulla natura cinetica microscopica del moto browniano veniva dalla osservazione che i dati sperimentali e le teorie cinetiche consentivano di valutare che (essendo le particelle in sospensione colloidale centinaia di milioni di volte più pesanti delle molecole del liquido solvente (acqua)) la velocità acquisita in ogni urto fosse così piccola da essere inosservabile ($\approx 2\,\mu$/sec). E poichè in un secondo il numero di urti era enorme e le variazioni di velocità in ognuno erano di segno casuale pareva inconcepibile che si potesse osservare un effetto medio non nullo. Argomento fallace, come ad esempio mise in rilievo Poincarè nel 1904 (senza però passare ad una teoria quantitativa), che anche notò (come altri) che l' ipotesi che il moto delle particelle colloidali fosse di natura cinetica era contraria alla termodinamica (si veda in seguito).

La fallacia di questo argomento era comunque ben nota, da tempo, come risulta ad esempio da un esoterico articolo che L. Bachelier (del 1900) pubblicò sull' austero Annales della Scuola Normale di Parigi, a poche pagine di distanza dalla traduzione francese dei Fondamenti della Geometria di Hilbert: è un articolo sulla teoria della speculazione [*in Borsa*], in cui viene analizzato proprio questo punto (cfr commenti alla fine del §2 seguente).

La teoria di Einstein prende le mosse dalla osservazione che le particelle in sospensione, anche se di taglia molto piu' grande di quella delle molecole del liquido solvente, possono essere trattate come se fossero molecole di grande massa e si può quindi applicare ad esse la meccanica statistica: per modo che esse vengono ad esercitare una pressione osmotica, proprio come le ordinarie soluzioni, verificante quindi la legge di Van t'Hoff (almeno a piccole concentrazioni). La legge diVan t'Hoff è dunque valida non solo per soluzioni di particelle microscopiche, ma anche per il calcolo della pressione parziale di particelle di arbitraria dimensione (ad es. biglie).

Questa idea era rivoluzionaria e, come si rese conto anche Einstein, con–

traria alla termodinamica classica ma non alla meccanica statistica e all' ipotesi atomica. E quindi si pose immediatamente il problema di trovare conseguenze macroscopiche osservabili.

Il moto delle particelle è dunque attribuito alle collisioni casuali con le molecole. E quindi è un moto aleatorio, almeno se osservato su scale di tempo τ grandi rispetto al tempo necessario a dissipare la velocità acquisita in un singolo urto. La dissipazione di tale velocità può essere stimata, per particelle macroscopiche, osservando che in un singolo urto con una molecola la velocità v acquisita è dissipata per attrito (esso stesso dovuto microscopicamente agli urti con le molecole del fluido) che esercita una forza F che secondo la legge di Stokes è:

$$m\frac{\partial v}{\partial t} \equiv F = -6\pi\eta R v \tag{1.1}$$

ove η è il coefficiente di viscosità del fluido e R il raggio delle particelle in sospensione e v la velocità; e, quindi, la scala di tempo caratteristica per la perdita della velocità acquisita in un singolo urto è $t_0 = 1/6\pi\eta m^{-1}R$. Questo è un tempo molto breve (ad es. se $R = 1\,\mu$, e se m si calcola supponendo la stessa densità del solvente, che si suppone acqua per cui $\eta = 10^{-3}\,$cgs, si trova che tale scala di tempo è $t_0 \approx 10^{-7}\,$sec).

Dunque su scale di tempo $\tau \gg t_0$ il moto avrà carattere diffusivo. Il moto diffusivo è un moto in cui si ha trasporto di materia solo in presenza di gradiente di densità. Si immagina allora che le particelle in sospensione abbiano una densità ν: questa è una situazione ideale che nulla ha necessariamente a che vedere con esperimenti sul moto browniano in cui la densità delle particelle è spesso praticamente nulla (ossia si osserva una singola particella che nel suo moto non interagisce con le altre). Sempre eseguendo questo esperimento ideale, si immagina di mantenere un gradiente di densità (delle sole particelle in sospensione ma *non* del solvente) stazionario $\partial_x\nu$ nella direzione x: allora per l' ipotesi che il moto sia diffusivo si avrà che il flusso Φ di particelle (numero Φ di particelle che passano la superficie unitaria nell' unità di tempo) è:

$$\Phi = D\partial_x\nu \tag{1.2}$$

dove D è il *coefficiente di diffusione*.

Ma questo flusso può essere calcolato anche da un punto di vista microscopico. Ossia si può osservare che il fatto che si mantiene un gradiente di densità implica (per la legge sulla pressione osmotica per le particelle macroscopiche) che deve essere presente un gradiente di pressione. Infatti se p è la pressione osmotica (ossia la pressione parziale delle particelle) si ha: $p = k_B T\nu$, essendo T la temperatura del fluido e k_B la costante di Boltzmann. Dunque $\partial_x p = k_B T\partial_x\nu$. E il gradiente di pressione genera una

forza F su ogni particella (dovuta alle collisioni con le altre particelle simili, presenti in questo esperimento ideale, ma assenti di solito negli esperimenti reali). La forza è, in media, ovviamente tale che $F\nu = \partial_x p = k_B T \partial_x \nu$; e quindi, per la legge di Stokes e su scale di tempo $\tau \gg t_0$, genera un moto uniforme di velocità $v = F/6\pi\eta R$ e flusso:

$$\Phi = \nu v = \nu \frac{F}{6\pi\eta R} = \frac{k_B T}{6\pi\eta R} \partial_x \nu \tag{1.3}$$

Questa seconda espressione del flusso deve coincidere con la precedente, ed eguagliandole si trova la *relazione di Einstein Smoluchowsky*:

$$D = \frac{k_B T}{6\pi\eta R} \tag{1.4}$$

(che dovrebbe essere interamente attribuita ad Einstein, si veda §2 seguente). La grandezza D è anche direttamente legata al valore medio (su varie traiettorie) $\langle \underline{r}(t)^2 \rangle$ del quadrato dello spostamento $\underline{r}(t)$ subito dalla particella colloidale nel tempo t: vedremo infatti che $\langle \underline{r}(t)^2 \rangle = 6Dt$. Poichè la grandezza di $\langle \underline{r}(t)^2 \rangle$ è misurabile direttamente al microscopio, si ha una prima relazione teorica che può essere confrontata con l' esperienza.

Prima di procedere conviene riassumere il filo logico della precedente analisi. L' uso di un esperimento ideale, metodo caratteristico del suo pensiero, permette ad Einstein di legare quantità microscopiche a quantità macroscopiche. La presenza del solvente qui fissa la temperatura e la scala di tempo sulla quale una particella si muove di moto diffusivo: per trovare poi il coefficiente di diffusione di una singola particella si immagina di avere un gas di particelle di densità arbitraria (ma piccola) ν (sebbene negli esperimenti classici le particelle in sospensione siano spesso così poche da potersi pensare isolate). Dunque nel primo passo dell' analisi si sostituisce la particella colloidale con un gas di particelle colloidali di densità arbitraria. Poi si immagina che su questo gas (che è una finzione dell' esperimento ideale) agisca una qualche causa esterna, pure essa parte dell' invenzione e non influenzante il fluido, e che, in situazione stazionaria, genera un gradiente di pressione e quindi (per la estensione alle particelle macroscopiche della legge della pressione osmotica) di densità. Se si puo' immaginare che il liquido in questione si comporti come un liquido di Navier Stokes si può allora calcolare a mezzo della formula di Stokes la velocità alla quale si muovono le particelle, legandola alla viscosità del fluido (strettamente parlando qui si deve supporre che le particelle siano macroscopiche) e quindi si può calcolare il flusso generato dal gradiente di pressione. Infine ci si ricorda che il moto delle particelle, viste individualmente, deve essere un moto diffusivo e quindi il flusso deve essere proporzionale al gradiente di densità. Si trovano così le relazioni che conducono alle (1.2),(1.3). E alla fine sono sparite tutte

le quantità ausiliarie usate per montare l'esperimento ideale e si conclude che se è valida la teoria cinetica allora una particella macroscopica (anche da sola) posta in un fluido e con esso in equilibrio termico deve avere un moto diffusivo legato alla viscosità dalla (1.4).

L' idea che le particelle macroscopiche si comportino come molecole (generando una pressione osmotica secondo la legge di Van t'Hoff, come le vere soluzioni), è una idea importante che fu essa stessa una delle novità del lavoro di Einstein: essa permise di far vedere anche a chi non aveva ancora creduto alla ipotesi atomica, che le leggi della termodinamica debbono essere di natura statistica (almeno dopo che Perrin mostrò sperimentalmente la correttezza della teoria sul moto browniano). Invero, poichè di solito si ritiene possibile costruire pareti semipermeabili a oggetti macroscopici, diviene possibile costruire dei cicli a temperatura costante in cui si usa la pressione osmotica e si converte calore in lavoro (perchè tali pareti sono una realizzazione di una entità simile al demonietto di Maxwell). Infatti si realizza un cilindro pieno di fluido e una parete semipermeabile che lo divide in due, ponendo nella parte di sinistra una soluzione colloidale che vede il lato impermeabile della parete e lasciando vuota (di particelle colloidali la parte di destra). Allora si può spostare il pistone verso destra facendo lavorare la pressione osmotica (a spese del bagno termico che mantiene il tutto a T costante). Si riporta poi la parete semipermeabile al centro del cilindro rovesciandola e si attende pazientemente che il colloide ritorni nella metà di sinistra, grazie alle collisioni casuali con la parete. A questo punto si può riprendere il ciclo: contro il principio di Carnot. L' occhio infinitamente sottile del demonietto di Maxwell può essere sostituito dal nostro microscopio: come commentò Poincaré (avendo in mente una diversa costruzione, facilmente immaginabile, pure conducente alla violazione del secondo principio). Perrin mise bene in luce questo aspetto commentando anche che la costruzione di una macchina funzionante come detto sopra avrebbe necessità di tempi inimmaginabilmente lunghi per estrarre quantità apprezzabili di energia. Però una analisi accurata della effettiva possibilità di costruire pareti semipermeabili a oggetti tipo i colloidi urta contro gravi difficoltà e, in realtà, ci si può convincere che tali oggetti sono eccessive idealizzazioni e non si può realizzare un reale moto perpetuo di seconda specie utilizzando le proprietà del moto browniano: si può invero osservare una violazione del principio di Carnot di cui parla Perrin, ma al variare del tempo tali violazioni vengono in media annullate. Fra le varie analisi di questo punto citiamo fra tutte quella che appare nelle lezioni di R. Feynman, vol. I, §46, ove la parete semipermeabile è sostituita da una ruota dentata con un meccanismo ad ancora che le consente di ruotare in un solo verso sotto gli urti delle particelle colloidali.

Il moto diffusivo produce uno spostamento r su un tempo t tale che il valor medio del suo quadrato è $\langle r^2 \rangle = 6Dt$, perche' la probabilità $f(\underline{x}, t)d^3\underline{x}$ di trovare una particella, inizialmente nell' origine, nel cubetto dx attorno alla posizione $\underline{x}$ è soluzione dell' equazione di diffusione $\partial_t f(\underline{x}, t) = D\Delta f(\underline{x}, t)$, e cioè:

$$f(\underline{x}, t) = \frac{e^{-\underline{x}^2/4Dt} d^3\underline{x}}{(4\pi Dt)^{3/2}} \tag{1.5.0}$$

(equazione che Einstein ricava imitando il metodo usato da Boltzmann per la deduzione dell' equazione di Boltzmann, trovando anche una espressione microscopica della D, della quale però non fa ulteriore uso). Il calcolo del valore medio del quadrato dello spostamento è allora semplicemente:

$$\langle \underline{r}(t)^2 \rangle = \int \underline{x}^2 f(\underline{x}, t)\, d^3\underline{x} = 6D\, t \tag{1.5.1}$$

Si vede dunque che, pur essendo ogni singolo urto tale da produrre una minima variazione di velocità subito quasi compensata da altre variazioni dello stesso segno e di segno opposto, tuttavia la particella subisce un moto che su un tempo lungo produce su ogni coordinata una variazione dell' ordine di $\sqrt{2Dt}$ (e $\sqrt{6Dt}$ se si osserva la variazione della posizione 3-dimensionale, ossia il vettore spostamento) che non solo non è nulla ma può essere considerevole e osservabile.

Come applicazione Einstein dedusse (1906) dalla misura della diffusione dello zucchero in sospensione in acqua il valore della k_B, quindi del numero di Avogadro N_A, trovando il risultato di $N_A = 4.7\, 10^{23}$, l'errore essendo dovuto principalmente alla incertezza dei dati sperimentali. Perrin trovò più tardi con esperimenti accurati e usando la teoria di Einstein un valore essenzialmente uguale a quello oggi accettato per N_A.

Einstein pervenne alla teoria del moto browniano senza veramente essere al corrente degli esperimenti in corso da circa 80 anni. Ma procedette per via deduttiva, basandosi su esperimenti ideali, a partire dall' osservazione che particelle anche macroscopiche, *dovevano* obbedire alle leggi della meccanica statistica e, in particolare, all' equipartizione dell' energia, che implicano per esse una pressione osmotica data dalla legge dei gas perfetti (legge di Van t'Hoff). E nel lavoro di Einstein appare anche chiaro che la legge del moto browniano conduce a moti che, se osservati su scale di tempo lunghe rispetto a t_0 (cfr (1.1)) ossia lunghe rispetto al microsecondo (come è necessariamente il caso), devono essere moti per i quali la velocità *non può essere definita* perchè gli spostamenti devono avere ordine di grandezza proporzionale a $\sqrt{t}$ invece che a t: ossia la velocità dipenderebbe dall' intervallo di tempo su cui la si misura e divergerebbe nel limite $t \to 0$ (o diverrebbe estremamente grande e fluttuante in valore e segno, all' avvicinarsi di t alla scala t_0 al di la della quale la teoria è inapplicabile).

La teoria del moto browniano fornì dunque un esempio di realizzazione in fisica di quelle curiosità matematiche che erano state, fino ad allora, le curve continue ma non differenziabili scoperte nell' '800 dai matematici che avevano rigorizzato il calcolo infinitesimale: lo stesso Perrin mise questo punto bene in luce.

L' ipotesi che la resistenza del fluido al moto della particella macroscopica segua la legge di Stokes non è in alcun modo fondamentale e questo è un aspetto caratteristico della teoria einsteiniana che, come vedremo, la distingue dalla teoria di Smoluchovski. Per cui se la ipotesi venisse mutata in $v = C_R F$ con C_R un' opportuna funzione del raggio R, la (1.4) verrebbe sostituita da $D = kTC_R$. Se, ad esempio, la particella si trovasse sospesa in un gas rarefatto, anzichè in un liquido incomprimibile, allora C_R sarebbe diverso.

Precisamente si vede che se la particella colloidale procede con una velocità v nel gas di densità ρ, allora il numero di particelle del gas che collidono con essa, per unità di tempo, con velocità (media) $-v_m$ è $\pi R^2(v + v_m)\rho/2$, mentre $\pi R^2(v - v_m)\rho$ è il numero di particelle che collidono con velocità (media) $+v_m$: le prime subiscono una variazione di impulso per unità di tempo pari a $2m(v_m + v)$ e le seconde pari a $2m(v_m - v)$. Quindi la forza esercitata dal fluido sulla particella è:

$$\frac{1}{2}\pi R^2[(v + v_m)^2 - (v_m - v)^2)2m = cR^2 v_m m\rho v \tag{1.5}$$

ove $c = 4\pi$. In questo calcolo si è supposto che metà delle particelle avesssero velocità uguale alla velocità assoluta media e metà velocità opposta e inoltre si è trattata la particella come un disco di raggio R perpendicolare alla direzione del moto.

La corretta trattazione dovrebbe assumere una distribuzione maxwelliana delle velocità e la forma sferica della particella. La valutazione delle correzioni non offre particolari difficoltà, nell' ipotesi che il gas sia assai rarefatto da poter trascurare i fenomeni di ricollisione (ossia di collisioni ripetute fra la particella e una stessa molecola di gas), e conduce al risultato finale identico alla (1.5) ma con un fattore diverso che sostituisce il 4π. E in definitiva si troverebbe:

$$D = \frac{kT}{cR^2 m v_m \rho} = \frac{kT}{cR^2\rho\sqrt{2mkT}} = \frac{\sqrt{kT}}{cR^2\rho\sqrt{2m}} \tag{1.6}$$

e la costante c risulta $2\sqrt{\pi}$. Per ovvii motivi il regime in cui vale questa forma della resistenza di attrito è detto regime "Doppler" ed è pertinente ai gas rarefatti, mentre il regime "Stokes", in cui vale la (1.1), riguarda la resistenza nei liquidi.

§2 La teoria di Smoluchowski.

La teoria di Smoluchowski, di poco successiva a quella di Einstein, getta luce sulle ipotesi di Einstein conducenti alla (1.5.0) attraverso un procedimento che implica ipotesi tacite del tipo di quelle sul caos molecolare implicite nella deduzione dell' equazione di Boltzmann; secondo la (1.5.0) il moto delle particelle colloidali deve essere un moto diffusivo, con spostamento quadratico medio proporzionale al tempo t.

Smoluchowski, per confutarel' argomento di Nägeli, considera un modello concreto microscopico delle collisioni in cui una particella, di massa M, è soggetta ad un gran numero di urti con le molecole, di massa m, del fluido ($\approx 10^{16}\,\mathrm{sec}^{-1}$ in molti casi).

Se $\underline{v}_k$ è la velocità della particella dopo k urti e se il k-mo urto avviene con una molecola che ha una velocità $\underline{v}$ prima dell' urto, si deduce dalle leggi dell' urto che:

$$\underline{v}_{k+1} \approx \underline{v}_k + \frac{m}{M}\mathcal{R}\underline{v} \qquad \text{se } |\underline{v}| \gg |\underline{v}_k| \tag{2.1}$$

ove $\mathcal{R}$ è una rotazione casuale (che dipende dal parametro d' urto della collisione, esso stesso casuale). Così, denotando le velocità medie v_m e v_M rispettivamente, la variazione di velocità induce una deviazione di un angolo ε dell' ordine di grandezza:

$$\varepsilon = \frac{\langle|\Delta\underline{v}_k|\rangle}{v_M} \approx \frac{m}{M}\frac{v_m}{v_M} = \sqrt{\frac{m}{M}} \tag{2.2}$$

poichè $mv_m^2 = Mv_M^2$ per l' equipartizione dell' energia. Si vede che l' angolo di deviazione è molto piccolo; e se τ è il tempo medio fra due collisioni, ossia:

$$\tau\pi R^2 v_m = 1 \tag{2.3}$$

lo spazio percorso in un tempo t durante il quale si hanno t/τ collisioni è:

$$\underline{r} = \sum_{k=1}^{n} \tau\underline{v}_k \tag{2.4}$$

Se ad ogni collisione si immagina che l' angolo fra le velocità iniziale e finale sia ε e orientato a caso sul cono di apertura ε rispetto alla velocità iniziale, è facile calcolare la distribuzione di probabilità della posizione $\underline{r}$; si trova che, per t grande, è una gaussiana con dispersione:

$$\langle\frac{\underline{r}^2}{t}\rangle = 6a\frac{kT}{cR^2m\rho v_m} \tag{2.5}$$

ove a è una opportuna costante numerica e c è introdotto in (1.6); e si vede che si ritrova sostanzialmente la formula di Einstein nel caso in cui la resistenza di Stokes è sostituita dalla resistenza Doppler di un gas rarefatto, come appunto ci si deve aspettare di trovare vista la natura del modello (a parte il fattore a).

Da questo Smoluchowski, con un salto logico abbastanza audace, deduce che se invece il fluido è un liquido incomprimibile allora si deve semplicemente sostituire, nel denominatore della (2.5) la resistenza in regime Doppler con quella in regime Stokes e trova allora:

$$D_{smol.} = aD_{einst.} \tag{2.6}$$

Si vede dunque che la teoria di Smoluchovskiè in un certo senso più ambiziosa di quella eisteiniana perchè vuole *dimostrare* che il moto del colloide è diffusivo senza trascurare del tutto le correlazioni temporali fra urti successivi (che Einstein, come si è già detto, trascura implicitamente); il modello che viene proposto è un modello in cui il fluido è pensato come un gas rarefatto e quindi non verifica la legge dell' attrito di Stokes. Strettamente parlando però il modello di Smoluchovski riguarda il caso, poco interessante per gli esperimenti del tempo, in cui il colloide è realizzato in un gas rarefatto e non è applicabile al caso di un fluido. Il metodo di Einstein è invece più generale e si applica ad entrambi i casi, pur non fornendo una reale giustificazione microscopica del carattere diffusivo del moto. Da un punto di vista concettuale Smoluchovski *non poteva ottenere* la formula di Einstein in quanto non era in grado di fornire un modello microscopico ragionevole di fluido in regime di Stokes (che ancora oggi sfugge all' indagine teorica): il suo metodo invero non è molto *oggettivo* neppure nel caso del gas rarefatto perchè conduce al risultato su D con un errore di un fattore a rispetto al risultato di Einstein: questo fattore è attribuibile alla crudezza delle approssimazioni (soprattutto la non chiarissima distinzione fra velocità e velocità media nel corso della deduzione della (2.5)) che certo non consentono di calcolare il valore di a correttamente. E tuttavia Smoluchowski non avendo il punto di vista macroscopico di Einstein è costretto a prendere sul serio il fattore a che trova, e trasferirlo (con il salto logico notato) al risultato errato nel caso del moto in un liquido.

Come nel caso del fattore c in (1.6) è possibile una teoria più precisa delle collisioni fra molecole e colloide, in cui si rimpiazzano i valori medi della velocità con i valori fluttuanti con la appropriata distribuzione maxwelliana: e così facendo si troverebbe $a = 1$. Se Smoluchovski avesse proceduto così, pur trovando il risultato esatto nel regime Doppler, avrebbe sempre dovuto fare un salto logico per trattare il caso di un colloide in un liquido.

Da quanto detto consegue che la teoria di Smoluchovsky *non era confrontabile* con l' esperienza, per ragioni intrinseche, e questo forse spiega

perchè Smolochovsky, come egli stesso dichiara, non abbia pubblicato i suoi risultati (che aveva a disposizione da vari anni) prima del lavoro di Einstein. Non è impossibile che la lettura della memoria di Einstein gli consentì il salto logico discusso sopra e necessario per il confronto con gli esperimenti (e che occupa poche righe della sua lunga memoria).

Successivamente Smoluchovski abbandonò il fattore a e adottò il valore 1 di Einstein.

Resta, però, che il lavoro di Smoluchovski è una pietra miliare nella teoria cinetica e la sua fu tra le prime di una serie di teorie che miravano a dedurre equazioni di continui macroscopici dissipativi a partire da modelli microscopici mirando ad ottenere le equazioni macroscopiche come descriventi teorie microscopiche su scale di tempo assai lunghe rispetto ai tempi microscopici in modo che il numero di eventi microscopici, fosse talmente grande da poter essere trattato usando i teoremi limite del calcolo delle probabilità (o tecniche equivalenti). L' uso del calcolo delle probabilità è la caratteristica innovatrice di queste teorie: già Lagrange, nella sua teoria della corda vibrante, aveva immaginato la corda come composta da tanti piccoli oscillatori, ma la sua deduzione era interamente "deterministica" tanto da apparire come un artificio.

Nel 1900, sei anni prima del lavoro di Smoluchovski, apparve una ricerca di L. Bachelier. La sua memoria, dal titolo assai poco invitante, sarebbe secondo alcuni rimasta inosservata, e oscurata dal lavoro di Einstein del 1905, ma presenterebbe la prima teoria del moto browniano. In realtà è possibile vedere una connessione fra la teoria della speculazione in borsa e la teoria del moto browniano solo *a posteriori*: tuttavia la memoria di Bachelier può forse essere considerata il primo lavoro in cui si ottengono, rigorosamente, equazioni macroscopiche dissipative a partire da modelli microscopici. Nel lavoro di Bachelier il moto browniano non viene menzionato e il suo modello di evoluzione dei listini dei prezzi è quello di un avanzamento o arretramento casuale di Δx in un tempo Δt con uguale probabilità: la novità rispetto alla classica teoria degli errori è che vi si considera il limite in cui Δx e Δt vengono fatti tendere a zero studiando la distribuzione delle variabili casuali associate *alle somme parziali* delle variazioni dei prezzi (nella teoria degli errori si studia solo la somma totale degli errori e l' indice di somma non ha l' interpretazione di tempo, bensì di indice enumeratore delle varie cause di errore). Si mostra che la distribuzione di probabilità del valore del listino al tempo t verifica una equazione di diffusione; si mostra inoltre che la distribuzione di probabilità di incrementi successivi è il prodotto di distribuzioni gaussiane indipendenti e si perviene ad una versione preliminare del processo stocastico gaussiano che sarà studiato in seguito da Wiener (in cui manca in sostanza "solo" la discussione della continuità delle traiettorie). Dunque si può considerare

questo lavoro come simile alla teoria di Smoluchovski, e quindi ben poco legato alla teoria di Einstein e nella sua analisi manca qualsiasi riferimento alla Fisica e alla termodinamica; ma abbiamo visto nella discussione precedente che è proprio qui che si trova una delle principali difficoltà della teoria del moto browniano.

§3 La teoria di Uhlenbeck Ornstein.

Come già detto, Einstein stesso (e anche Smoluchovski) notarono che la teoria del moto browniano era valida per osservazioni spaziate nel tempo di una quantità minima grande rispetto alla scala di tempo caratteristica per la perdita della velocità acquisita in un singolo urto che è $t_0 = 1/6\pi\eta m^{-1}R$. Per tempi piccoli rispetto a questa scala di tempo ha ancora senso definire la velocità delle particelle e il moto non può essere descritto dal processo diffusivo caratteristico del moto browniano propriamente detto, osservato su scale di tempo $\gg t_0$. Le traiettorie del moto browniano osservate su scale di tempo grandi rispetto a t_0 appaiono erratiche, irregolari e se si tenta di calcolare la velocità dividendo lo spazio percorso per il tempo corrispondente si trova un risultato che dipende dalla grandezza dell' intervallo di tempo e che diviene sempre più grande man mano che si accorcia l' intervallo di tempo fra le osservazioni. Questo riflette immediatamente il fatto che, su queste scale di tempo, la media del valore assoluto dello spostamento è proporzionale a $\sqrt{t}$, invece che a t. Però questa "divergenza" della velocità cessa appena si esamina il moto delle particelle su scala di tempo piccola rispetto a t_0.

Si pone allora il problema di svolgere una teoria del moto che lo descriva sia nella fase "normale", a tempi piccoli, sia nella fase "browniana" a tempi grandi. Langevin propose un modello matematico assai semplice per le equazioni complete del moto browniano. Immaginò che le successive collisioni con le molecole del fluido potesssero essere descritte, per ogni componente cartesiana del moto, da una forza impulsiva casuale $F(t)$ e quindi l' equazione del moto di una coordinata di una particella di colloide fosse:

$$m\dot{v} = -\lambda v + F(t) \tag{3.1}$$

ove λ è il coefficiente di attrito per il moto della particella colloidale nel fluido (ossia $6\pi\eta R$ nel caso di un fluido in regime Stokes, e $cR^2 v_m m\rho$ nel caso di regime Doppler, cfr (1.5)).

L' equazione di Langevin può essere discussa una volta assegnata una legge che descriva le proprietà della forza casuale F. Il modello proposto da Uhlenbeck e Ornstein fu che $F(t)$ fosse un *rumore bianco*, ossia fosse tale che:

1) non ci fosse alcuna correlazione fra i valori di $F(t)$ ai vari istanti e,

2) che la legge di distribuzione di una n-pla $F(t_1), F(t_2), \ldots, F(t_n)$ di valori della forza, osservati a n istanti qualunque $t_1 < t_2 < \ldots < t_n$, fosse descritta da una distribuzione gaussiana, e
3) che il valore medio di $F(t)$ fosse identicamente nullo al variare di t.

Questo da un lato conduce alla nozione di processo stocastico gaussiano (centrato), e alla più generale nozione di processo stocastico, e dall' altro conduce ad una teoria del moto browniano come processo stocastico "esattamente risolubile".

Si consideri un processo stocastico (ossia una distribuzine di probabilità su uno spazio di eventi rappresentabili come funzioni $F(t)$ di una variabile (o più)) che sia *gaussiano* a media nulla, ossia tale che la probabilità di osservare una n-pla $F(t_1)$, $F(t_2)$,...,$F(t_n)$ di valori della forza misurando la $F(t)$ a n istanti qualunque $t_1 < t_2 < \ldots < t_n$, è una distribuzione gaussiana a media nulla. Si può dimostrare che un tale processo (ossia la distribuzione di probabilità delle scelte delle funzioni $t \to F(t)$) è univocamente determinato dalla funzione di correlazione a due punti (detta *covarianza* o *propagatore*). Tale funzione è definita come valore medio del prodotto delle funzioni a due istanti arbitrarii t_1, t_2:

$$C(t_1, t_2) = \langle F(t_1) F(t_2) \rangle \tag{3.3}$$

e ciò vuol dire che la distibuzione gaussiana della probabilità di una n-pla qualunque di valori di F osservati a n istanti diversi è calcolabile in termini della C (e, precisamente, è elementarmente esprimibile in termini della matrice inversa della matrice $C(t_i, t_j)$ $(i, j = 1, 2, \ldots n)$.

In questo linguaggio il "rumore bianco" è definito come il processo gaussiano con covarianza:

$$C(t, t') = f^2 \delta(t - t') \tag{3.4}$$

ove f^2 è una costante e δ è la funzione di Dirac.

La soluzione esplicita della (3.1) è possibile se C è data da (3.4) e risulta, come dimostrarono Uhlenbeck e Ornstein, che ciascuna componente sia della velocità che della posizione delle particelle verificanti la (3.1),(3.4), con dati iniziali s_0 per la posizione e v_0 per la velocità, sono processi gaussiani a media non nulla. Se $\beta = \lambda/m \equiv t_0^{-1}$, la loro media è, all' istante t:

$$\begin{aligned} \overline{s}(t) =& s_0 + \frac{v_0}{\beta[1 - e^{-\beta t}]} \\ \overline{v}(t) =& v_0 e^{-\beta t} \end{aligned} \tag{3.5}$$

e la distribuzione di probabilità di una componente v della velocità al tempo t è una gaussiana:

$$G(v, t) = \Big(\frac{m}{2\pi k T (1 - e^{-2\beta t})} \Big)^{1/2} \exp\{ - \frac{m}{2kT} \frac{(v - \overline{v}(t))^2}{(1 - e^{-2\beta t)})} \} \tag{3.6}$$

mentre la distribuzione di una componente $s(t)$ della posizione è una gaussiana con centro in $\overline{s}(t)$ e con dispersione quadratica:

$$\sigma(t) = \frac{kT}{m\beta^2}(2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}) \xrightarrow[t\to\infty]{} \frac{2kT}{m\beta} t \equiv 2D_{einst.}t \qquad (3.7)$$

ossia:

$$H(s,t) = \Big(\frac{1}{2\pi\sigma(t)}\Big)^{1/2} e^{-(s-\overline{s}(t))^2/2\sigma(t)} \qquad (3.8)$$

Le formule ora descritte si riducono alle precedenti della teoria di Einstein nel limite $t \to \infty$, ma hanno valore anche se $t < t_0$ e quindi risolvono il problema della teoria del moto delle particelle colloidali su scale di tempo dell' ordine di t_0 o inferiori.

Uhlenbeck e Ornstein calcolarono anche le distribuzioni di probabilità "congiunte" dei valori $v(t_1), s(t_1), \ldots, v(t_n), s(t_n)$ per $t_1, \ldots t_n$ arbitrari e il processo gaussiano risultante (cioè la distribuzione di probabilità delle funzioni a due componenti $t \to (v(t), s(t))$) è pertanto chiamato *processo di OrnsteinUhlenbeck*.

§4 *La teoria di Wiener.*

Dal punto di vista matematico ci si può porre il problema di come appaia il moto di un punto tale che, a partire da ogni istante, la probabilità di subire uno spostamento di $\underline{r}$ nel tempo t è data dalla distribuzione asintotica (per $t \to \infty$ o $t \gg t_0$) del moto browniano, ossia da:

$$P(\underline{r}, t) = \frac{e^{-\underline{r}^2/4Dt}}{(4\pi Dt)^{3/2}} \qquad (4.1)$$

Evidentemente si troverà un moto con la notevole proprietà di *non* avere velocità definita in *nessun istante*: proprio come il moto browniano osservato su scale di tempo lunghe rispetto al tempo t_0 introdotto precedentemente.

In realtà il fatto stesso che un tale oggetto sia definibile è notevole e il contributo di N. Wiener (1923) fu proprio quello di dimostrare che il processo gaussiano con probabilità di transizione (4.1) (già introdotto da Bachelier, nel citato articolo) è ben definito dal punto di vista matematico e che *con probabilità* 1 le traiettorie descritte dal processo in questione sono *continue*, e anzi Hölder continue con esponente α (con $\alpha < 1/2$), con probabilità 1.

Il processo gaussiano che descrive la probabilità delle traiettorie $t \to \underline{r}(t)$ in cui gli incrementi di $\underline{r}$ sono distributiti con legge gaussiana (4.1) è detto *processo di Wiener* o, semplicemente, nella teoria delle probabilità *moto*

browniano: e dal punto di vista fisico corrisponde alla descrizione del comportamento asintotico del moto di una particella colloidale in un fluido, per tempi grandi rispetto al tempo caratteristico t_0 di rilassamento (mentre per tempi piccoli o generici tale moto è piuttosto descritto dal processo di Ornstein-Uhlenbeck). Qui lo chiameremo *processo di Wiener* per distinguerlo dal moto browniano nel senso fisico della parola (che è un moto che genera traiettorie dotate di velocità su scale di tempo abbastanza piccole, a differenza delle traiettorie del processo di Wiener).

Il processo di Wiener in quanto processo stocastico, e nonostante si tratti di una astrazione matematica, ha un grande interesse e appare nei più svariati campi della Fisica e della Matematica. La prima applicazione fu di dare luogo a varie formule *di quadratura* che consentono di esprimere le soluzioni di varie equazioni differenziali alle derivate parziali in forma *esplicita* come integrali su famiglie di curve distribuite a caso con la legge che corrisponde a quella del processo di Wiener.

Ovvviamente il calcolo di questi integrali non è di solito più semplice che la risoluzione delle equazioni stesse con metodi più tradizionali. Tuttavia la natura *esplicita* delle formule ne permette l' uso per una rappresentazione intuitiva delle soluzioni di certe equazioni differenziali e anche al fine di ottenere stime a priori delle soluzioni.

Un esempio classico è la teoria dell' equazione del calore:

$$\begin{aligned} \partial_t u =& D\,\Delta u \\ u|_{t=0} & u_0 \end{aligned} \tag{4.2}$$

la cui soluzione può essere scritta come:

$$u(\underline{x},t) = \int d\underline{y} \int P(d\underline{\omega})^t_{\underline{x},\underline{y}} u_0(\underline{y}) \tag{4.3}$$

ove l' integrale è esteso a tutte le curve continue $\tau \to \underline{\omega}(\tau)$ che per $\tau = 0$ partono dal punto $\underline{y}$ e al tempo t giungono in $\underline{x}$, e sono distribute con la distribuzione del processo di Wiener, condizionata ad avere punto di arrivo $\underline{x}$.

La interpretazione della (4.2) è la seguente: *il calore è dotato di moto browniano*, ossia si trasferisce eseguendo un moto descritto dal processo di Wiener. Pertanto la quantità di calore $u(\underline{x},t)$ che al tempo t si trova in $\underline{x}$ è ottenibile immaginando che la quantità di calore inizialmente in un punto generico $\underline{y}$ si distribuisca equamente fra tutte le traiettorie del processo di Wiener che escono da $\underline{y}$ e così la quantità di calore che si trova in $\underline{x}$ al tempo t (ossia $u(\underline{x},t)$) sia la somma su tutte le traiettorie browniane che pervengono in $\underline{x}$ dopo un tempo t, ciascuna apportando una quantità di calore proporzionale alla quantità di calore $u_0(\underline{y})d\underline{y}$ che si trova inizialmente

nel punto ove hanno origine; e il fattore di proporzionalità è proprio uguale alla frazione di traiettorie browniane che partono da $d\underline{y}$ e giungono in $\underline{x}$ nel tempo t.

La formula (4.3) è il caso più semplice di una classe di formule *risolutive* di equazioni differenziali alle derivate parziali; ad esempio l' equazione:

$$\begin{aligned} \partial_t u =& D\,\Delta u + V(x)u \\ u(\underline{x},0) =& u_0(\underline{x}) \end{aligned} \tag{4.4}$$

può essere risolta *esplicitamente* da:

$$u(\underline{x},t) = \int d\underline{y} \int P^t_{\underline{x},\underline{y}}(d\underline{\omega}) e^{-\int_0^t V(\underline{\omega}(\tau))d\tau} u_0(\underline{y}) \tag{4.5}$$

che è chiamata *formula di Feynman Kac.*

Questa ultima formula ha varie estensioni, importanti nella Matematica e nella Fisica moderna, in campi assai diversi: dalla teoria delle probabilità e delle equazioni differenziali alla Mecccanica statistica, alla teoria dei campi e quantizzati e perfino ai fondamenti della teoria dei quanti (si veda in proposito l' analisi delle " variabili nascoste" nella teoria di E. Nelson o in quella di D. Bohm). Si può dire che nelle teorie menzionate le formule di soluzione *esplicita* derivate dagli esempi appena dati giocano spesso un ruolo simile a quello giocato dalle *quadrature* nella meccanica classica.

Il processo di Wiener ha avuto, ed ha, una grande importanza nella teoria delle probabilità ove ha introdotto una grande quantità di idee, problemi e soluzioni a problemi preesistenti.

Come esempio di proprietà matematiche del processo di Wiener citiamo le seguenti.

1) La *legge di regolarità* di P. Levy: risponde alla questione di quale sia il comportamento della variazione $\omega(t) - \omega(t')$ di una componente $\omega(t)$ di una traiettoria $t \to \underline{\omega}(t)$ del processo di Wiener osservata in due istanti t, t' vicini, e contenuti in un intervallo $[0, \bar{t}]$ prefissato. Si è già detto che, come mostrato da Wiener, la traiettoria ha la proprietà di essere Hölder continua con esponente $\alpha < 1/2$, arbitrario. Questo significa che, comunque si fissi $\alpha < 1/2$ si avrà *con probabilità* 1:

$$\lim_{t-t' \to 0} \frac{|\omega(t) - \omega(t')|}{|t - t'|^\alpha} = 0 \tag{4.6}$$

se $0 \le t, t' \le \bar{t}$. La arbitrarietà di $\alpha < 1/2$ rende interessante la domanda di quale sia il valore ottimale per α, se esistente. La legge di Levy dice che un α ottimale non esiste, ma fornisce in un certo senso la risposta a quale sia la regolarità di ω, perchè dice:

$$\limsup_{\substack{t-t' \to 0 \\ 0 \le t,t' \le \bar{t}}} \frac{|\omega(t) - \omega(t')|}{(4D|t - t'| \log \frac{1}{|t-t'|})^{1/2}} = 1 \tag{4.7}$$

con probabilità 1.

2) Ma la legge di regolarità di Levy non da informazioni sulle proprietà della traiettoria nell' intorno di un dato punto: invero la (4.7) da solo il *comportamento peggiore*, ossia da solo la misura della massima irregolarità. Se si fissa l' attenzione su un dato istante $\bar{t}$ in generale la traiettoria sarà meno irregolare, con probabilità 1. Questo infatti è il contenuto della *legge locale del logaritmo iterato* di V. Kintchin: tale legge dá la proprietà di regolarità di una traiettoria ad un istante prefissato $\bar{t}$; fissando $\bar{t} = 0$ dice:

$$\limsup_{t\to 0} \frac{|\omega(t)|}{(4Dt\log(\log\frac{1}{t})^{1/2}} = 1 \tag{4.8}$$

con probabilità 1. La (4.8) non è incompatibile con la (4.7): anzi dice che il comportamento peggiore descritto dalla (4.7) è violato con probabilità 1 ad un istante *prefissato*!

3) Le due leggi precedenti riguardano il comportamento delle traiettorie a tempi finiti; ci si può domandare quale sia il comportamento a tempi lunghi di una data traiettoria. La teoria di Einstein Smoluchovski prevede infatti che il moto si allontani dall' origine di una distanza che *in media* è crescente proporzionalmente a $\sqrt{t}$. Un' analisi della discussione di Einstein e Smoluchovski indica che per *media* si debba intendere la media su un insieme statistico di moti. Se, all' istante t, si misura il quadrato variazione di una coordinata rispetto al valore nella posizione iniziale, per molte particelle browniane si trova, in media, $2Dt$. Ma questo non significa che se si fissa un solo moto $t \to \underline{\omega}(t)$, cioè una sola particella, e lo si osserva al variare di t allora ogni sua coordinata $\omega(t)$ differisce dal valore iniziale di una quantità che, elevata al quadrato, cresce al più come $2Dt$ nel senso che il limite superiore per $t \to \infty$ di $\omega(t)^2/t$ è $2D$. Infatti la reale crescita è data dalla *legge globale del logaritmo iterato* (di Kintchin):

$$\limsup_{t\to+\infty} \frac{|\omega(t)|}{\Big(4Dt\log(\log t)\Big)^{1/2}} = 1 \tag{4.9}$$

con probabilità 1.

Poichè la (4.9) riguarda le proprietà del processo di Wiener per tempi grandi essa esprime proprietà importanti anche per le osservazioni fisiche del moto browniano; *a differenza* delle due precedenti leggi che riguardano proprietà a tempi brevi, e quindi sono proprietà caratteristiche del processo di Wiener ma non dei moti browniani reali (che piuttosto sono descritti dal processo di Ornstein Uhlenbeck. Tuttavia dal punto di vista fisico è difficilissimo effettuare esperimenti così precisi da osservare una correzione allo spostamento proporzionale alla radice di un logaritmo iterato.

4) La (4.9) non invalida la misurabilità di D basata sulla osservazione di una singola traiettoria. Tali misure si effettuano osservando lo spostamento $\omega(\tau)$ di una coordinata al variare di τ fra 0 e t. Si pone poi:

$$X(t) = t^{-1} \int_0^t \omega(\tau)^2 d\tau \tag{4.10}$$

e si tenta di interpolare i dati $X(t)$ confrontandoli con la funzione $2Dt$. Il procedimento è corretto, almeno asintoticamente per $t \to \infty$, perchè si dimostra che vale:

$$\lim_{t\to\infty} \frac{X(t)}{2Dt} = 1 \tag{4.11}$$

con probabilità 1. È questo il *teorema ergodico* per il processo di Wiener.

5) Le curve descritte dal processo di Wiener sono piuttosto *irregolari*, come esplicitamente mostrano le leggi di Levy e di Kintchin. Ci si può domandare quale sia la dimensione dell' insieme descritto dalla traiettoria. Se la dimensione dello spazio è ≥ 2 *la dimensione della traiettoria è* 2, nel senso della teoria della dimensione di Hausdorff. Questo significa essenzialmente che se si vuole ricoprire la traiettoria con sferette di raggio $\leq 1/n$ ne sono necessarie circa n^2 (nel senso che, per $n \to \infty$, $n^{2-\varepsilon}$ non sono sufficienti se $\varepsilon > 0$ mentre $n^{2+\varepsilon}$ sono invece sufficienti).

Ma aver dimensione 2 può essere espresso in altri modi intuitivamente equivalenti, ma non esattamente equivalenti, all' aver dimensione di Hausdorff 2. E la discussione di tali nozioni alternative di dimensione illustrano aspetti sottili della struttura delle traiettorie del processo di Wiener. Ad esempio se si considerano due punti diversi in R^d e da ciascuno si fanno partire due traiettorie che evolvono come prescritto dal processso di Wiener si trova che tali traiettorie hanno un punto in comune con probabilità 1 se $d = 2, 3$, mentre non hanno punti in comune se $d \geq 4$: se le traiettorie fossero *ordinarie* superfici bidimensionali dovrebbero *genericamente* intersecarsi per $d = 2, 3$ e anche 4.

6) Un' altra celebre proprietà del processo di Wiener, dovuta a Wiener stesso, fa vedere interessanti legami con l' analisi armonica e la serie di Fourier. Si consideri una successione $g_0, g_1, \ldots$ di variabili aleatorie gaussiane equidistribuite ed indipendenti le une dalle altre e la distribuzione di ciascuna di esse sia $(2\pi)^{-1/2} \exp -g^2/2$.

Si pone:

$$\omega(t) = \frac{t}{\pi^{1/2}} g_0 + \sum_{n\geq 1} \sum_{k=2^{n-1}}^{2^n-1} \left(\frac{2}{\pi}\right)^{1/2} \frac{\sin kt}{k} g_k \tag{4.10.1}$$

e la funzione aleatoria $\omega(t)$, per $0 \leq t \leq \pi$, ha la distribuzione di probabilità (indotta da quella supposta per i coefficienti g_k) identica a quella del processo di Wiener (con $D = 1/2$) a dimensione 1.

In conclusione il processo di Wiener è un' astrazione matematica generata dal moto browniano e ne descrive il comportamento a tempi grandi: tuttavia è una entità matematica di grande interesse che fornisce applicazioni ai più svariati campi della Matematica e pure della Fisica.

Bibilografia

Brush, S.: *The kind of motion that we call heat*, North Holland, 1976, vol. II, C. 15.
Pais, A.: *Sottile è il Signore*, Boringhieri, 1990.
Bachelier,, L.: *Théorie de la spéculation*, Annales Scientifiques de l' École Normale Superieure, 17, 21-36, 1900.
Einstein, A.: *Investigations on the theory of the brownian movement*, con note di R. Fürth e A. Cowper, ed. Dover, 1956
Perrin, J.: *Les atomes*, Gallimard, 1948.
Feynman, R.: *Lezioni di Fisica.*, vol I, §41,&§46, Zanichelli, 1966.
Wax, N.:*Selected papers on noise and random stochastic processes*, Dover, 1954
Ito, K., McKean, H.: *Diffusion processes and their sample paths*, Springer Verlag, 1965
Nelson, E.: *Dynamical theories of brownian motion*, Princeton U. Press, 1967.

æ

Entropia e informazione.

voce per il DIZIONARIO DELLE SCIENZE FISICHE
Autore della voce Giovanni Gallavotti
versione 1: agosto 1985

§1 *Introduzione. Statistica di una successione di simboli.*
§2 *Complessità o informazione di successioni. Entropia dei sistemi dinamici.*
§3 *Significato della complessità. Teorema di Shannon McMillan.*
§4 *Entropia nei sistemi dinamici ed entropia termodinamica.*
§5 *La complessità dei moti in Meccanica Statistica.*
§6 *Alcune applicazioni alla Teoria dell'Informazione.*

§1. Introduzione. Statistica di una successione di simboli.

Il problema di dare una misura quantitativa della "informazione" o " complessità" di una successione di numeri (o " segnale", " messaggio", "testo", "configurazione") è molto naturale.

Si consideri una successione $\underline{\sigma} = (\ldots \sigma_{-1}, \sigma_0, \sigma_1, \sigma_2, \ldots)$ prodotta in un modo qualunque.

Si possono tener presenti i seguenti casi:

1) successioni di risultati di misure eseguite, in vari istanti regolarmente spaziati, su di un evento che si produce in un esperimento;

2) successioni di numeri generate da un calcolatore a mezzo di un dato programma di calcolo;

3) successione delle lettere dell' alfabeto in un testo scritto;

4) gli stati microscopici degli spin degli atomi di una sostanza ferromagnetica cristallina misurati al variare del sito cristallino i lungo una retta tracciata nel cristallo.

L' ultimo esempio rende chiaro che si può anche essere interessati alla nozione di complessità di successioni $\overline{\sigma} = (\sigma_\xi)_{\xi \in L}$ ove L è un reticolo, o sottoinsieme di un reticolo, non necessariamente unidimensionale: ad esempio nel caso 4) anziché considerare gli spin lungo una retta nel cristallo si potrebbero considerare tutti gli spin del cristallo indicizzati dai punti ξ del cristallo stesso che definisce il reticolo L.

Ci occuperemo principalmente di successioni "unilatere" (indicizzate cioè da $i = 1, 2, 3, \ldots$ nel caso unidimensionale o dai punti ξ di coordinate positive nel caso di successioni a più dimensioni) e *infinite*: quest'ultima caratteristica è richiesta perché le nozioni di statistica, ergodicità, complessità, entropia etc. associate alle successioni sono nozioni "asintotiche".

Questo non vuol dire che non sia poi necessario stabilire analoghe nozioni per successioni finite abbastanza lunghe (o, a più dimensioni, "grandi"): la trattazione del caso idealizzato infinito è tuttavia preliminare alla trattazione del caso finito che è assai difficile e più interessante. Ma qui ci limiteremo, per esigenze di spazio, al caso ideale, infinito.

In teoria dell' informazione uno dei motivi per cui ci si interessa alla misura della complessità di una successione ("testo") è quello di cercare il più corto codice che contenga la stessa informazione del testo dato.

Data una successione $\overline{\sigma}$ la prima fondamentale nozione da associare ad essa è quella della sua "statistica". La statistica di $\underline{\sigma}$ è la collezione delle frequenze di apparizioni in $\underline{\sigma}$ di tratti $\overline{\underline{\tau}} = (\overline{\tau}_1, \ldots, \overline{\tau}_p)$ di lunghezza finita: matematicamente la frequenza di $\overline{\underline{\tau}}$ in $\underline{\sigma}$ è definita contando il numero $\mathcal{N}_N(\overline{\tau}_1, \ldots, \overline{\tau}_p | \underline{\sigma})$ di valori di $j \leq N$ per cui:

$$\sigma_{j+1} = \overline{\tau}_1, \sigma_{j+2} = \overline{\tau}_2, \ldots, \sigma_{j+p} = \overline{\tau}_p \tag{1.1}$$

e quindi ponendo:

$$\nu(\overline{\sigma}_1, \ldots, \overline{\sigma}_p|\underline{\sigma}) = \lim_{N\to\infty} \mathcal{N}_N(\overline{\sigma}_1, \ldots, \overline{\sigma}_p|\underline{\sigma})/N \tag{1.2}$$

se il limite esiste. La (1.2) è la "frequenza" con cui $\underline{\overline{\sigma}} = (\overline{\sigma}_1, \ldots, \overline{\sigma}_p)$ appare in $\underline{\sigma}$.

Se $\underline{\sigma}$ è una successione a più dimensioni la definizione è analoga: un tratto finito $\underline{\overline{\sigma}}$ sarà $(\overline{\sigma}_\xi)_{\xi\in\Lambda_p}$ ove Λ_p è un cubo di lato p consistente dei punti ξ a coordinate $\xi_i > 0$ e $\leq p$ e $\mathcal{N}_N(\underline{\overline{\sigma}}|\underline{\sigma})$ sarà il numero di ξ in Λ_N tali che $\sigma_{\xi+\eta} = \overline{\sigma}_\eta$ con η in Λ_p; in (1.2) ovviamente si dovrà ora dividere per il volume $|\Lambda_N| = N^d$ se d è la dimensione della successione, anziché per N.

Una successione ha una statistica se i limiti (1.2) esistono per ogni $\underline{\overline{\sigma}} = (\overline{\sigma}_1, \ldots, \overline{\sigma}_p)$: e la statistica di $\underline{\sigma}$ è la collezione delle frequenze definite in (1.2).

Se ad esempio $\underline{\sigma}$ è un lungo libro, idealmente infinito, scritto in italiano il tratto, o "parola", (gg) ha una certa frequenza di apparizione e così anche (ghg) (il primo ha frequenza positiva e il secondo, ovviamente, nulla).

Se $\underline{\sigma}$ è la successione dell' esempio 4) e $\sigma_j = \pm 1$ la parola (1) ha frequenza definita e uguale alla magnetizzazione media per sito del ferromagnete; se la magnetizzazione è positiva la parola (1,1) ha frequenza maggiore di $(-1, 1)$, etc... .

Empiricamente è possibile verificare che certe parole hanno frequenza definita solo se ci si limita alle parole più corte.

Però esistono teoremi generali per successioni generate attraverso precisi meccanismi; tali teoremi garantiscono, sotto ipotesi piuttosto generali, che le successioni in esame hanno frequenze definite.

Il tipico meccanismo che genera successioni è idealizzabile così.

Sia M uno spazio (si pensi a R^d ovvero ad una superficie senza frontiera o allo spazio $\{0, 1\}^Z$ di tutte le possibili successioni di zeri e uni) su cui è definita una trasformazione S che lascia invariante una distribuzione di probabilità μ su M:

$$\mu(E) = \mu(S^{-1}E) \tag{1.3}$$

ove E è un qualunque insieme (μ–misurabile).

Allora si può usare il "sistema dinamico" (M, S, μ) per generare successioni definendo una partizione $\mathcal{P} = \mathrm{P}_1, \ldots, \mathrm{P}_n)$ di M in insiemi disgiunti μ–misurabili e quindi definire $\underline{\sigma}$ estraendo a caso un punto di M con distribuzione μ e ponendo:

$$S^k x \in P_{\sigma_k} \tag{1.4}$$

cioè $\underline{\sigma}$ è la "storia" del moto di x osservata su $\mathcal{P}$.

È questo il metodo in base al quale si producono nei calcolatori successioni di "numeri casuali": si considera una trasformazione S: $[0, 1] \to [0, 1]$ e

"scelto a caso" un numero $x \in [0,1]$, "seme" della successione $\underline{\sigma}$, si definisce:

$$\begin{aligned} \sigma_k &= 0 \quad \text{se} \quad S^k x \in [0, 1/2) \\ \sigma_k &= 1 \quad \text{se} \quad S^k x \in [1/2, 1] \end{aligned} \tag{1.5}$$

essendo qui $\mathcal{P} = ([0,1/2),[1/2,1])$. Se S è "opportuna" si generano successioni di 0 e 1 che hanno la stessa statistica delle successioni prodotte dai lanci successivi di un dado a due facce (si veda la voce "Numeri Casuali").

Analogamente si possono generare successioni a più dimensioni usando sistemi dinamici d–dimensionali (M,S,μ) ove ora $S = (S_1,\dots,S_d)$ sono d trasformazioni di M in se commutanti fra loro: se $\xi = (\xi_1,\dots,\xi_d) \in Z^d$ si pone $S^\xi = S^{\xi_1}\dots S^{\xi_d}$ e data una partizione $\mathcal{P} = (\mathrm{P}_1,\dots,\mathrm{P}_n)$ di M si definisce $(\sigma_\xi)_{\xi \in \dot{Z}^d}$ ponendo $S^\xi x \in \mathrm{P}_{\sigma_\xi}$ se x è un punto di M scelto a caso (e $\underline{\sigma}$ è la "storia" di x su $\mathcal{P}$).

In generale se $\underline{\sigma} = \underline{\sigma}(x)$ è la storia di un punto x, preso a caso in M con distribuzione μ e osservato su una partizione $\mathcal{P}$ del sistema dinamico (M,S,μ) (a una o più dimensioni) si dimostra, "teorema di Birkhoff", che $\underline{\sigma}$ *ha frequenze definite con* μ*–probabilità* 1 *sulle scelte di* x, cioè ha statistica.

Inoltre si dimostra che la statistica di $\underline{\sigma}(x)$ è, *con probabilità* 1 *sulle scelte di* x, "ergodica": cioè date due parole $\overline{\sigma} = (\overline{\sigma}_1,\dots,\overline{\sigma}_{\overline{p}})$ e $\hat{\underline{\sigma}} = (\hat{\sigma}_1,\dots,\hat{\sigma}_{\hat{p}})$ la frequenza con cui appare la parola $\overline{\underline{\sigma}}$ seguita dopo r siti dalla parola $\hat{\underline{\sigma}}$ (evento che denotiamo $\overline{\underline{\sigma}}\tau^r\hat{\underline{\sigma}}$) è in media uguale al prodotto delle frequenze di $\overline{\underline{\sigma}}$ e $\hat{\underline{\sigma}}$:

$$\lim_{N\to\infty} N^{-1} \sum_{r=0}^{N-1} \nu(\overline{\underline{\sigma}}\,\tau^r\,\hat{\underline{\sigma}}|\underline{\sigma}) = \nu(\overline{\underline{\sigma}}|\underline{\sigma})\nu(\hat{\underline{\sigma}}|\underline{\sigma}) \tag{1.6}$$

ossia parole diverse sono in media distribuite indipendentemente sulla successione $\underline{\sigma}$.

Ad esempio se $\underline{\sigma}$ è una successione di numeri 0 e 1 prodotti indipendentemente a caso con probabilità q e $(1-q)$, ogni parola di lunghezza p contenente α lettere 0 e $(p-\alpha)$ lettere 1 ha frequenza uguale alla sua probabilità, e cioè $q^\alpha(1-q)^{p-\alpha} > 0$.

Per futuro riferimento è utile ricordare che esiste anche una nozione equivalente a quella di ergodicità per un sistema dinamico (M,S,μ) formulabile senza riferimento a successioni o partizioni ad esso associate.

Si dice che il sistema dinamico (M,S,μ) è "ergodico" se, con μ–probabilità 1 sulle scelte di x, la frequenza di visita ad un prefissato insieme (μ–misurabile) E è uguale alla probabilità $\mu(E)$ di E:

$$\lim_{N\to\infty} N^{-1} \sum_{j=0}^{N-1} \chi_E(S^j x) = \mu(E) \tag{1.7}$$

ove χ_E è la funzione caratteristica di E.

È facile vedere che se (M, S, μ) è ergodico e $\mathcal{P} = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$ è una partizione di M, le storie $\underline{\sigma}(x)$ dei punti x di M hanno, *con μ–probabilità* 1 *sulle scelte di* x, la stessa statistica indipendente da x ed ergodica e per ogni $\underline{\overline{\sigma}} = (\overline{\sigma}_1, \ldots, \overline{\sigma}_p)$ finita si ha

$$\nu(\underline{\overline{\sigma}}|\underline{\sigma})) = \mu\left(\bigcap_{j=1}^{p} S^{-j}\mathrm{P}_{\sigma_j}\right) . \tag{1.8}$$

Tutti i concetti precedenti concernenti l' ergodicità si estendono in modo naturale ai casi di sistemi dinamici a più dimensioni ma tralasceremo questa discussione.

§2. *Complessità o informazione di successioni. Entropia di sistemi dinamici.*

Si può definire la complessità di una successione $\underline{\sigma}$ di simboli $\sigma_i \in (1, \ldots, n)$ a statistica definita ed ergodica (che è il caso "generico", si veda §1) in vari modi.

Un primo modo conduce alla nozione di "complessità assoluta" ed è, come vedremo, piuttosto insoddisfacente e di scarsa utilità.

Si conta il numero $N(p|\underline{\sigma})$ di parole di lunghezza p che hanno frequenza positiva in $\underline{\sigma}$; tale numero è a priori limitato da n^p, evidentemente, e se si pone:

$$s_0(\underline{\sigma}) = \lim_{p\to\infty} \frac{1}{p} \log N(p|\underline{\sigma}) \tag{2.1}$$

e si può dimostrare facilmente che il limite esiste sempre. La (2.1) è la misura della "complessità assoluta" di $\underline{\sigma}$.

Per successioni a d dimensioni la definizione è analoga e le parole di lunghezza p sono sostituite dalle successioni $\overline{\sigma}$ associate al cubo Λ_p di lato: $\underline{\overline{\sigma}} = (\sigma_\xi)_{\xi\in\Lambda_p}$ e $1/p$ diviene $1/|\Lambda_p| \equiv 1/p^d$.

La quantità $s_0(\underline{\sigma})$ misura il numero di parole che hanno frequenza positiva. Ad esempio se $\underline{\sigma}$ è una successione di 0 e 1 prodotti estraendo 0 e 1 a caso con probabilità q e $(1-q)$, allora ognuna delle 2^p parole possibili di lunghezza p con α zeri e $(p-\alpha)$ uni ha frequenza e probabilità uguali a $q^\alpha(1-q)^{p-\alpha} > 0$; dunque:

$$s_0(\underline{\sigma}) = \lim_{p\to\infty} \frac{1}{p} \log 2^p = \log 2 \tag{2.2}$$

La complessità assoluta è insoddisfacente come è ben messo in luce dalla (2.2): infatti predice uguale complessità per successioni prodotte con $q \ll 1$ e con $q = 1/2$, che sono in realtà molto diverse.

Intuitivamente le prime sono semplici in quanto costituite in massima parte di 1 e le seconde sono più complesse in quanto costituite di uguali quantità di 0 e 1.

La seguente definizione di Shannon è assai più soddisfacente ed interessante perché distingue il ruolo delle parole più probabili da quello delle altre nel computo della complessità.

Data una successione $\underline{\sigma}$ ergodica ed a statistica definita e dato $\varepsilon > 0$ si dividono le parole di lunghezza p in due classi C_p^0 e C_p^1 in modo che la frequenza *totale* di apparizione delle parole della classe C_p^0 sia minore di ε:

$$\sum_{\overline{\sigma} \in C_p^0} \nu(\overline{\underline{\sigma}}|\underline{\sigma}) < \varepsilon \ . \tag{2.3}$$

È chiaro che C_p^0 può essere scelto in vari modi (ad esempio vuoto!): però siamo interessati a vedere quanto piccola può essere C_p^1 se C_p^0 verifica (2.3).

Consideriamo dunque una scelta di C_P^0, che chiameremo $\overline{C}_P^0$, per cui C_P^1 contiene il minimo numero di parole: sia $N(p|\underline{\sigma}, \varepsilon)$ questo numero. Si pone poi:

$$s(\underline{\sigma}) = \lim_{\varepsilon \to 0} \lim_{p \to \infty} \frac{1}{p} \log N(p|\sigma, \varepsilon) \tag{2.4}$$

che verrà chiamata "complessità" o "informazione" di $\underline{\sigma}$: se $\underline{\sigma}$ ha statistica definita si può dimostrare che i limiti (2.4) esistono.

Poiché $\lim_{\varepsilon \to 0} N(p|\underline{\sigma}, \varepsilon) = N(p|\underline{\sigma}) =$ numero di parole a frequenza definita e positiva, cfr. (2.1), si ha:

$$s_0(\underline{\sigma}) = \lim_{p \to \infty} \lim_{\varepsilon \to \infty} (1/p) \log N(p|\underline{\sigma}, \varepsilon) \ . \tag{2.5}$$

È questo però uno dei casi in cui i limiti non si possono permutare! e $s \le s_0$.

Ad esempio, come un non difficile e molto istruttivo calcolo mostra, nel caso delle successioni di 0 e 1 estratti indipendentemente a caso con probabilità q e $(1-q)$ si ha (cfr. (2.2)):

$$s_0(\underline{\sigma}) = \log 2 \ , \quad s(\underline{\sigma}) = -q \log q - (1-q) \log(1-q) \ . \tag{2.6}$$

Si vede quindi che $s(\underline{\sigma}) \to 0$ se $q \to 0$ o 1 e cioè la complessità di $\underline{\sigma}$ è piccola se uno dei due simboli $0, 1$ ha probabilità piccola, mentre è massima se hanno uguale probabilità ($q = 1/2, s = s_0 = \log 2$).

L' estensione delle due nozioni di complessità al caso di successioni a più dimensioni è semplice e coinvolge solo semplici cambiamenti di notazione ($\overline{\underline{\sigma}} = (\overline{\sigma}_1, \ldots, \overline{\sigma}_p)$ diviene $\overline{\underline{\sigma}} = (\sigma_\xi)_{\xi \in \Lambda_p}$ e $1/p$ diviene $1/|\Lambda_p| \equiv 1/p^d$, etc).

È possibile introdurre anche una interessante nozione di complessità di un sistema dinamico (M, S, μ) che è strettamente legata a quella, appena definita, delle successioni.

Semplicemente si considera una generica partizione $\mathcal{P} = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$ di M in n insiemi μ–misurabili a due a due disgiunti e per ogni $x \in M$ si genera la sua storia $\underline{\sigma}(x)$ su $\mathcal{P}$ (si veda §1, (1.4)). Con probabilità 1 la $\underline{\sigma}(x)$ ha statistica ergodica e si pone:

$$\begin{aligned} c(\mu) &= \sup_{\mathcal{P}} \quad \operatorname*{esssup}_x \; s(\underline{\sigma}(x)) \\ s(\mu) &= \sup_{\mathcal{P}} \quad \int \mu(\mathrm{d}x) s(\underline{\sigma}(x)) \end{aligned} \tag{2.7}$$

dove $\operatorname{esssup} \cdot \equiv \inf_{A,\mu(A)=0} \sup_{x \notin A} \cdot$ è l' estremo superiore essenziale.

La quantità $c(\mu)$ è l' "entropia massima" di (M, S, μ) mentre $s(\mu)$ è l' " entropia" (media) o "invariante di Kolmogorov–Sinai". I primi risultati che si trovano su queste quantità sono i seguenti.

Si dimostra che se (M, S, μ) è ergodico:

$$c(\mu) = s(\mu) \ . \tag{2.8}$$

Si dimostra inoltre che se P è "generante", ossia se le storie su P di x e y sono tali che $\underline{\sigma}(x) = \underline{\sigma}(y)$ implica $x = y$, si ha:

$$\begin{aligned} c(\mu) &= \operatorname*{esssup} \; s(\underline{\sigma}(x)) \\ s(\mu) &= \int \mu(\mathrm{d}s) s(\underline{\sigma}(x)) \ . \end{aligned} \tag{2.7}$$

Dunque l' entropia è calcolabile senza considerare estremi superiori sulle partizioni e basta calcolarla usando una partizione generante (se esiste), (teorema di Sinai).

Si può dimostrare dunque che le partizioni che generano i moti più complessi sono quelle generanti e raffinando ulteriormente una partizione generante, cioè aumentando la precisione delle singole osservazioni, *non si aumenta l' informazione che si ottiene.* Ovviamente aumentando la precisione aumenta l' informazione che si ottiene da una *singola* misura ma quanto dice il teorema appena citato è che se si eseguono infinite misure in successione temporale può non essere necessario usare una grande precisione in ogni singola misura. Nel §3 si vedrà un altro aspetto di questa stessa questione.

Inoltre se (M, S, μ) è ergodica e $\mathcal{P}$ è una partizione generante di M, si ha con μ–probabilità 1 sulle scelte di x:

$$c(\mu) = s(\mu) = s(\underline{\sigma}(x)) \ . \tag{2.8}$$

Infine si dimostra che se $P = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$ è una partizione generante di M si ha la seguente formula per l' entropia:

$$s(\mu) = -\lim_{p\to\infty} \frac{1}{p} \sum_{\overline{\sigma}_1,\ldots,\overline{\sigma}_p} \mu(\bigcap_{j=1}^{p} S^{-j} P_{\overline{\sigma}_j}) \log \mu(\bigcap_{j=1}^{p} S^{-j} P_{\overline{\sigma}_j}) \tag{2.9}$$

e si osservi che $\bigcap_{j=1}^{p} S^{-j} P_{\overline{\sigma}_j}$ è l' insieme dei punti x la cui storia su P, $\underline{\sigma}(x)$, inizia con $\overline{\sigma}_1, \ldots, \overline{\sigma}_p$.

Se P non è generante il membro di destra della (2.9) definisce una quantità $s(\mu, P)$ tale che:

$$s(\mu) = \sup_P s(\mu, P) \tag{2.10}$$

Tutte le precedenti nozioni e risultati si estendono ai sistemi dinamici a più dimensioni.

Si vedrà come, pur non essendo la stessa nozione, la complessità di una successione e l' entropia di un sistema dinamico siano intimamente legate e sostanzialmente coincidenti.

§3. Significato della complessità. Il teorema di Shannon–McMillan.

Sia $\underline{\sigma}$ una successione dotata di statistica ergodica (che è il caso "generico", cfr. §1, (1.6)).

Il seguente teorema di struttura di Shannon–McMillan chiarisce la natura e l' interesse della nozione di complessità o entropia:

i) la complessità della successione $\underline{\sigma}$ è calcolabile dalle frequenze di apparizione (cfr. §1).

$$s(\underline{\sigma}) = \lim_{p\to\infty} -(1/p) \sum_{\overline{\sigma}_1,\ldots,\overline{\sigma}_p} \nu(\overline{\sigma}_1, \ldots, \overline{\sigma}_p|\underline{\sigma}) \log \nu(\overline{\sigma}_1, \ldots, \overline{\sigma}_p|\underline{\sigma}) \; ; \tag{3.1}$$

ii) dato $\varepsilon > 0$ è possibile dividere le parole di lunghezza p in due classi $C^{\text{rare}}(p)$ e $C^{\text{frequenti}}(p)$, in modo che:

$$\sum_{(\overline{\sigma}_1,\ldots,\overline{\sigma}_p)} \nu(\overline{\sigma}_1, \ldots, \overline{\sigma}_p|\underline{\sigma}) < \varepsilon \tag{3.2}$$

cioè la frequenza delle parole rare è *in totale* inferiore ad ε, mentre se $\underline{\sigma} \in C^{\text{frequenti}}(p)$ si ha:

$$e^{-(s(\sigma)+\varepsilon)p} < \nu(\overline{\sigma}|\underline{\sigma}) < e^{-(s(\sigma)-\varepsilon)p} \tag{3.3}$$

e, quindi, il numero di elementi di $C^{\text{frequenti}}(p)$ è:

$$e^{(s(\sigma)-\varepsilon)p} < \#\, C^{\text{frequenti}}(p) < e^{(s(\sigma)+\varepsilon)p} \tag{3.4}$$

ove $\# A =$ numero degli elementi nell' insieme A.

Dunque la (3.4) è una formula alternativa per il calcolo della complessità e ii) ci dice che a parte un insieme di parole rare, di frequenza totale $< \varepsilon$, tutte le altre hanno frequenze comparabili nel senso (3.3) ed il loro numero è misurato dalla complessità $s(\underline{\sigma})$ di $\underline{\sigma}$.

Così se l' "alfabeto" in cui $\underline{\sigma}$ è scritta contiene n simboli, i "messaggi" di lunghezza p estratti da $\underline{\sigma}$ possono essere scritti ("salvo casi rari" di frequenza complessiva ε) con un alfabeto con $\tilde{n} < n$ simboli purché $\tilde{n}$ sia tale che $s(\underline{\sigma}) < \log \tilde{n}$.

Alternativamente si possono riscrivere le parole di lunghezza p (salvo casi rari come sopra) nello stesso alfabeto ma in modo che siano più corte e cioè di lunghezza $\tilde{p}$: la massima compressione è tale che $s(\underline{\sigma})p = \tilde{p} \log n$:

$$\tilde{p}/p = s(\underline{\sigma})/\log n\ . \tag{3.5}$$

Anche quanto detto sopra si estende in modo ovvio a successioni a più dimensioni.

Un'altra interessante proprietà, già notata al §2 e resa più quantitativa dal precedente teorema, della entropia e della complessità delle successioni generate a mezzo di osservazioni dei moti di un sistema dinamico (M, S, μ) con partizioni di M è che, contrariamente all' intuizione, la complessità dei moti *non può divenire infinitamente grande infittendo la partizione* $P = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$ sulla quale si vogliono osservare i moti di un sistema dinamico ergodico (per semplicità) dato, a meno che $s(\mu) = +\infty$: la massima complessità viene raggiunta non appena P è così fitta da essere generante. È dunque inutile eseguire osservazioni troppo precise sulla evoluzione di un sistema dinamico: infittendo P, cioè aumentando il numero n degli insiemi P_i e quindi la precisione delle osservazioni non è più utile se $n > \tilde{n}$, indipendentemente da n!

È pertanto importante elaborare criteri per dare stime a priori di $s(\underline{\sigma})$ per le successioni $\underline{\sigma}$ prodotte con dati meccanismi. Ad esempio ottenute a partire da un sistema dinamico (M, S, μ) osservandone i moti su una partizione $\mathcal{P}$ generante.

Ad esempio un caso notevole in cui si può a priori stabilire una stima su $s(\mu)$ è quando M è una superficie limitata priva di frontiera e S è un diffeomorfismo che lascia invariante una misura di probabilità di S equivalente alla misura di volume. In questo caso se λ è il massimo coefficiente di dilatazione di un elemento di linea $\underline{d}\, x$ su M:

$$\lambda = \sup \frac{|S\, \underline{d}\, x|}{|\, \underline{d}\, x|}, \tag{3.6}$$

si può dimostrare ("teorema di Kouchnirenko") che:

$$S(\mu) \leq d \log \lambda \tag{3.7}$$

che mette in luce la relazione che intercorre fra la complessità dei moti su una superficie e la "espansività" della trasformazione che li genera.

La (3.7) fa anche intravedere la possibilità, che di recente è stata effettivamente usata nella interpretazione di risultati di esperimenti sulla turbolenza, di stabilire una connessione quantitativa fra esponenti di Lyapunov entropia di un sistema dinamico e moti caotici.

Un risultato preciso in questo senso, molto utile nelle applicazioni, è la " formula di Pesin" della quale citiamo solo un caso particolare ma indicativo.

Se (M, S, μ) è un sistema dinamico ergodico e "iperbolico" (che è una delle nozioni matematiche precise che traducono il concetto di un sistema dinamico in cui i moti siano molto instabili e le traiettorie di punti vicini divergano, in genere, esponenzialmente) e se $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q > 0$ sono gli esponenti di Lyapunov (si veda la voce) non negativi di S rispetto a μ e se μ è una misura invariante proporzionale alla misura di volume su M (o più in generale assolutamente continua rispetto ad essa) allora:

$$s(\mu) = \sum_{j=1}^{q} -\lambda_i \log \lambda_i \tag{3.8}$$

che è più precisa, nel caso considerato, della (3.7).

§4. Entropia ed informazione nei sistemi dinamici ed entropia termodinamica.

Si consideri ora un sistema meccanico di N particelle di massa m racchiuse in un volume V a pareti riflettenti e interagenti via una forza di energia potenziale $\Phi(\underline{q}) = \sum \varphi(\underline{q}_i - \underline{q}_j)$ con $\varphi(\underline{r}) = +\infty$ se $|\underline{r}| < r_0$ ("cuore duro") e $|\varphi(\underline{r}) < |\underline{r}|^{-3-\alpha}$, $C, \alpha > 0$ se $|\underline{r}| > r_1$ ("temperatezza"), e $r_0, r_1 > 0$, si veda Insiemi Statistici, §, (2.17), (2.18).

L' esistenza del cuore duro è qui assunta per semplicità e molto di quanto segue si può estendere ai casi più generali. Anche la condizione di temperatezza può essere in certi casi resa più debole (si vedano le voci Stabilità della Materia, Insiemi Statistici, Meccanica Statistica Classica).

Si consideri ora la distribuzione canonica:

$$\mu^{(V)}(\mathrm{d}\underline{p}\,\underline{d}\,\underline{q}) = \frac{e^{-\beta\left(\sum \underline{p}_i^2/2m + \Phi(\underline{q})\right)}}{h^{3N} N!\, Z(\beta, V)}\, \underline{d}\,\underline{p}\,\underline{d}\,\underline{q} \tag{4.1}$$

(si veda Meccanica Statistica Classica e Insiemi Statistici, (1.7), (1.27) e (6.7)).

Facciamo vedere come si possa associare a (4.1) un sistema dinamico tridimensionale naturale.

Se V_0 è un volume cubico fissato in V, sia $\overline{\mu}^{(V)}_{V_0}$ la distribuzione di probabilità che descrive la probabilità di trovare in V_0 una data configurazione $\underline{q}_1, \dots, \underline{q}_g$ con $\underline{q}_i \in \underline{d}\,\underline{q}_i$. Tale $\overline{\mu}^{(V)}_{V_0}$ si ottiene fissando queste coordinate ed integrando $\mu^{(V)}$ su tutte le altre coordinate con il vincolo che le $\underline{q}_j$ con $j > g$ siano fuori di V_0.

Il risultato ha necessariamente la forma:

$$\overline{\mu}^{(V)}_{V_0}(\mathrm{d}q_1 \dots \underline{d}\,q_g) = f^{(V)}_{V_0}(q_1, \dots, q_g) \quad \underline{d}\,q_1 \dots \underline{d}\,q_g/g! \tag{4.2}$$

con $f^{(V)}_{V_0}$ opportuna: la barra su μ ricorda che si è eseguita l' integrazione anche sugli impulsi $\underline{p}_1, \dots, \underline{p}_g$ corrispondenti a $\underline{q}_1, \dots, \underline{q}_g$.

Tali impulsi sono poco interessanti in quanto la loro distribuzione è nota a priori e maxwelliana e solo la distribuzione delle $\underline{q}$ in V_0 è non banale.

Si considera poi il "limite termodinamico" $V \to \infty$, con $V/N = v$ e β fissi e si definisce in questo modo una distribuzione μ di probabilità sulle configurazioni $\underline{q}$ del sistema infinito: la μ–probabilità di trovare in V_0 g particelle in $d\underline{q}_1 \dots d\underline{q}_g$ è per definizione:

$$f_{V_0}(\underline{q}_1, \dots, \underline{q}_g)\frac{\underline{d}\,\underline{q}_1 \cdots \underline{d}\,\underline{q}_g}{g!} = \lim_{V\to\infty} f^{(V)}_{V_0}(\underline{q}_1, \dots, \underline{q}_g)\frac{\underline{d}\,\underline{q}_1 \cdots \underline{d}\,\underline{q}_g}{g!} \tag{4.3}$$

e supporremo che il limite in (4.3) esista, per semplicità, e sia invariante per traslazione:

$$f_{V_0+\xi}(\underline{q}_1 + \xi, \dots, \underline{q}_g + \xi) = f_{V_0}(\underline{q}_1, \dots, \underline{q}_g) \tag{4.4}$$

ove ξ è un vettore di R^3.

La misura μ che si ottiene è definita dalla famiglia delle $f_V(\underline{q})$ al variare di $\underline{q}$ in V_0 e prende il nome di "misura di Gibbs" (configurazionale) sullo spazio M delle successioni $\underline{q}$ tali che $|\underline{q}_i - \underline{q}_j| > r_0$ di potenziale $\Phi(\underline{q}_1, \dots, \underline{q}_g)$, densità v^{-1}, temperatura $T = 1/k\beta$ (k= costante di Boltzmann).

Più in generale si dicono "misure di Gibbs" tutte le distribuzioni che si possono ottenere, con procedimento simile a quello descritto nella (4.2), (4.3), a partire da un qualunque insieme statistico (microcanonico, canonico, gran canonico con o senza condizioni al contorno, si veda la voce Insiemi Statistici, §5). Considereremo solo quelle invarianti per traslazioni, cfr. (4.4).

Però sullo spazio M si possono definire molte altre distribuzioni μ di probabilità invarianti per traslazioni che in generale non sono stati di Gibbs (per

un potenziale dato Φ) per nessun valore della temperatura $T = 1/k\beta$ o della densità v^{-1}.

Ad ognuna di queste distribuzioni si può associare un sistema dinamico tridimensionale (M, S, μ) ove $S = (S_1, S_2, S_3)$ è la terna di trasformazioni che traslano di una unità di lunghezza nelle direzioni x, y o z di R^3.

E ognuna di queste distribuzioni μ può essere usata per calcolare la densità media delle particelle e l' energia potenziale specifica media associata da μ alle configurazioni se esse vengono pensate "interagire tramite l' energia potenziale Φ":

$$v(\mu)^{-1} = \lim_{V_0 \to \infty} \frac{1}{V_0} \sum_{g=0}^{\infty} \int f_{V_0}(\underline{q}_1, \dots, \underline{q}_g) g \, \frac{\underline{d}\underline{q}_1 \dots \underline{d}\underline{q}_g}{g!} \tag{4.5}$$

$$U_{\text{pot}}(\mu) = \lim_{V_0 \to \infty} \frac{v(\mu)}{V_0} \sum_{g=0}^{\infty} \int f_{V_0}(\underline{q}_1, \dots, \underline{q}_g) \Phi(\underline{q}_1, \dots, \underline{q}_g) \frac{\underline{d}\underline{q}_1 \dots \underline{d}\underline{q}_g}{g!}$$

Si può inoltre definire l' entropia termodinamica di μ:

$$s_{\text{term}}(\mu) = \tag{4.6}$$

$$= \lim_{V_0 \to \infty} \frac{-v(\mu)}{V_0} \sum_{g=0}^{\infty} \int f_{V_0}(\underline{q}_1, \dots, \underline{q}_g) \log f_{V_0}(\underline{q}_1, \dots, \underline{q}_g) \frac{\underline{d}\underline{q}_1 \dots \underline{d}\underline{q}_g}{g!}$$

Ci poniamo il problema di stabilire una connessione fra $s_{\text{term}}(\mu)$ e l' entropia del sistema dinamico (M, S, μ) almeno nel caso che μ sia uno stato di Gibbs con potenziale φ e parametri β, v.

Il punto chiave è qui un importante risultato generale di Meccanica Statistica. Si consideri la relazione (che va pensata come definizione dell quantità $f(\beta, v)$:

$$-\beta f(\beta, v) = \sup_{\mu} (s_{\text{term}}(\mu) - \beta U_{\text{pot}}(\mu) + \frac{3}{2} \log \beta^{-1}) \; . \tag{4.7}$$

Allora si può dimostrare che:

i) l' estremo superiore è un massimo che può essere raggiunto in una o più distribuzioni μ;

ii) ciascuna delle μ ove la funzione in (4.7) raggiunge il massimo è uno stato di Gibbs configurazionale per il potenziale Φ con densità v^{-1} e temperatura $T = 1/k\beta$. Se tali distribuzioni sono più d'una esse descrivono fasi diverse in equilibrio termodinamico, si veda Insiemi Statistici §4,5;

iii) il valore di $k(s_{\text{term}}(\mu) + \frac{3}{2} \log \beta^{-1})$ e di $U_{\text{pot}}(\mu)$ calcolate nei punti in cui il massimo in (4.7) vien raggiunto sono, a meno di una costante additiva l' entropia e l' energia interna, rispettivamente, per particella del modello di termodinamica statistica associato al potenziale φ a densità v^{-1} e temperatura $1/k\beta$.

Avendo trovato l' interpretazione termodinamica della (4.6) ci proponiamo di vedere la relazione fra (4.6) e l' entropia del sistema dinamico (M, S, μ) nei casi in cui μ è una distribuzione di Gibbs configurazionale, cioè una misura di probabilità sullo spazio delle fasi che descrive uno stato di equilibrio macroscopico (si veda Meccanica Statistica, §5 e Insiemi Statistici, §1,2).

Si vorrebbe, al fine di una piena giustificazione dei nomi che stiamo usando, che $s(\mu)$ ed $s_{\text{term}}(\mu)$ siano uguali o differiscano banalmente (ad esempio per una costante additiva).

Si vede però facilmente che, nei casi che consideriamo e se $s_{\text{term}}(\mu)$ è finita, allora $s(\mu) = +\infty$, con il seguente ragionamento euristico.

Costruiamo una partizione $\mathcal{P}$ dello spazio M: dividiamo lo spazio ambiente in cubetti ottenuti per traslazione di un cubetto di lato $\varepsilon \ll r_0$ centrato nell' origine: così nel cubo unitario ci saranno ε^{-3} cubetti. Si definisce una partizione $\mathcal{P}$ dello spazio M delle configurazioni $\underline{q}$ di infinite particelle in 2 insiemi: $P_0 =$ insieme delle configurazioni con nessuna particella nel cubetto centrato nell' origine e $P_1 =$ insieme delle configurazioni con una particella nel detto cubetto (questa è una partizione di M perchè c'è l' ipotesi di cuore duro).

Poi dividiamo le configurazioni $\underline{q} \in M$ in classi caratterizzate dai numeri di particelle che si trovano nei vari cubetti: cioè dagli ε^{-3} numeri (0 o 1) di particelle in ciascuno dei cubetti (tali numeri sono 0 o 1 perché c'è, per ipotesi, un cuor duro fra le particelle).

La storia $\underline{\sigma}$ di $\underline{q}$ su P rispetto all' azione delle traslazioni $S = (S_1, S_2, S_3)$ unitarie di $\underline{q}$ nelle tre direzioni coordinate altro non è che la specificazione del numero σ_ξ (0 o 1) di particelle che si trovano nel cubetto di lato ε e centro ξ (i centri dei cubetti si immaginano situati sui punti del reticolo di passo ε) al variare di ξ.

Si consideri ora il cubo $V_0 \equiv \Lambda_p$ di lato p: e siano $x_1, \ldots, x_g$ i centri dei $g = \varepsilon^{-3}$ cubetti di lato ε prima considerati contenuti in Λ_p vediamo che specificare la storia di $\underline{q} \in M$, rispetto alle traslazioni di S_ξ con vettori $\xi \in \Lambda_p$, significa specificare quali sono i g cubetti di lato ε, fra quelli che pavimentano V_0, ad essere occupati.

Se denotiamo la probabilità che i g cubetti occupati siano proprio quelli con centro in $x_1, \ldots, x_g$ con $f_{V_0}(x_1, \ldots, x_g)\varepsilon^{3g}$, ne segue che l' entropia $s(\mu, P)$ è:

$$\lim_{V_0 \to \infty} \frac{-v(\mu)}{V_0} \sum_{g=0}^{\infty} \sum_{x_1 \ldots x_g} \frac{f_{V_0}(x_1, \ldots, x_g)\varepsilon^{3g}}{/g!} \log(f_{V_0}(x_1, \ldots, x_g)\varepsilon^{3g}) \tag{4.8}$$

ove $g!$ è stato inserito per sommare sulle g–ple di punti, distinti due a due, $x_1, \dots, x_g$ senza tener conto dell' ordine.

Confrontando (4.8) con (4.6) si vede che:

$$s(\mu, P) = s_{\text{term}}(\mu) - \\ - \lim_{V_0 \to \infty} \Big[\frac{-v(\mu)}{V_0} \sum_{g \geq 0} \sum_{x_1 \dots x_g} g \frac{f_V(x_1, \dots, x_g) g \varepsilon^{3g})}{g!} \Big] \log \varepsilon^3 \qquad (4.9)$$

e allora si vede che $s(\mu) = +\infty$ (poiché per calcolare $s(\mu) = \sup s(\mu, P)$ si deve considerare il limite $\varepsilon \to 0$, cfr. (2.10), perché la partizione P non è generante per $\varepsilon > 0$).

Si vede però anche che $s(\mu)$ differisce a meno di una costante *infinita* da $s_{\text{term}}(\mu)$ e questa costante infinita è una costante numerica adimensionale (uguale a $\log \varepsilon^3$, per $\varepsilon \to 0$) perché, questioni di rigore matematico a parte, il termine in parentesi quadra in (4.9) tende a 1 per $V_0 \to \infty$, essendo il prodotto del volume specifico $v(\mu)$ per la densità media.

Resta dunque giustificato il nome di entropia usato in generale per $s(\mu)$ e si vede che se μ è una misura di Gibbs allora la sua entropia termodinamica è una misura della complessità della struttura spaziale delle configurazioni q tipiche di μ.

§5. *La complessità dei moti in Meccanica Statistica.*

Si è visto che l' entropia termodinamica classica è legata alla complessità spaziale delle configurazioni di equilibrio di un gas.

Però un sistema di N particelle racchiuso in un volume V a pareti riflettenti con energia potenziale $\Phi(\underline{q})$ può anche essere pensato come un sistema dinamico in modo profondamente diverso da quello visto al §4.

È infatti un sistema dinamico (W, S, μ) ove W è la superficie di energia totale costante e uguale ad U situata nello spazio delle fasi a $6N$ dimensioni, S è l' evoluzione temporale hamiltoniana per un tempo τ prefissato e μ è la misura di Liouville ristretta alla superficie di energia U.

Si può allora definire l' entropia media di questo sistema dinamico come $S_{\text{din}}(U, V)$ e poi definire $s_{\text{din}}(u, v) = V^{-1} S_{\text{din}}(U, V)$, con $u = U/V, v = V/N$.

Questa quantità è di difficile analisi e non si sa se esiste il limite:

$$\sigma(u, v) = \lim_{V \to \infty} s_{\text{din}}(u, v) \qquad (5.1)$$

quando u, v vengono mantenuti fissi ("limite termodinamico").

Né sembra che ci siano ragioni a priori per credere che ci sia una qualsiasi relazione fra $\sigma(u, v)$ e l' entropia termodinamica.

Probabilmente $\sigma(u, v)$, che misura la complessità temporale del moto e non la complessità spaziale, ha piuttosto a che fare con qualche proprietà di non equilibrio del sistema, quale la viscosità o qualche altro coefficiente di trasporto.

Ma il problema è interamente aperto.

§6. *Alcune applicazioni alla Teoria dell' Informazione.*

Menzioniamo qui alcune applicazioni della teoria dell' entropia dei sistemi dinamici, più direttamente collegate alla teoria dell' informazione.

Il risultato forse più naturale è la risposta ad una antica questione.

Dati due "schemi di Bernoulli" (si veda la voce) determinati dalla probabilità dei singoli eventi $(p_1, \ldots, p_n)$ e $(p'_1, \ldots, p'_{n'})$; date cioè due successioni $\underline{\sigma}$ e $\underline{\sigma}'$ costruite con n e n' simboli, rispettivamente, estratti a caso con probabilità rispettive $(p_1, \ldots, p_n)$ e $(p'_1, \ldots, p'_{n'})$ ci si domanda quando sia possibile stabilire un "codice", cioè una funzione che trasformi successioni $\underline{\sigma}$ in successioni $\underline{\sigma}' : \underline{\sigma} = C(\underline{\sigma})$ in modo che "ogni" successione $\underline{\sigma}$ distribuita con la statistica del primo schema di Bernoulli venga trasformata in una successione $\underline{\sigma}'$ distribuita con la statistica del secondo schema di Bernoulli e "viceversa": "ogni" vuol dire con probabilità 1, "viceversa" vuol dire che C è invertibile quasi ovunque e che C^{-1} ha la stessa proprietà.

Più in generale si possono sostituire i due schemi di Bernoulli con due processi di Markov sugli stessi spazi di successioni.

La risposta alla questione appena posta è estremamente semplice: il codice C esiste se e solo se i due sistemi dinamici hanno la stessa entropia. Nel caso degli schemi di Bernoulli questo significa se e solo se vale la:

$$\sum_{j=1}^{n} p_j \log p_j = \sum_{j=1}^{n'} p'_j \log p'_j \tag{6.1}$$

e per i processi di Markov si hanno formule analoghe.

Questo teorema di Ornstein, successivamente rafforzato in vari casi da una abbastanza efficace ed esplicita costruzione del codice C (M. Keane, M. Smorodinski) è, nella sua dimostrazione, una raffinata analisi della nozione di entropia e delle sue relazioni con la teoria dei codici.

L' entropia in generale non è però un "invariante completo" per i sistemi dinamici: ossia non basta in generale che due sistemi dinamici (M, S, μ) e (M', S', μ') abbiano la stessa entropia perché esista un codice C (definito ed invertibile a meno di insiemi di punti di probabilità complessiva nulla), $x' = C(x)$, che trasforma punti di M in punti di M' e viceversa, in modo che μ venga trasformata in μ' e S in S'.

È però vero che se (M, S, μ) è un sistema dinamico ergodico e $s(\mu) > 0$ si può in esso definire una partizione $\mathcal{P} = (\mathrm{P}_1, \ldots, \mathrm{P}_n)$, in generale fatta di insiemi estremamente complicati, tale che le storie $\underline{\sigma}(x)$ di μ–quasi tutti i punti di M abbiano la statistica di uno schema di Bernoulli di entropia uguale a $s(\mu)$.

Dunque i sistemi dinamici ergodici con entropia positiva hanno "in se" " moti caotici" e cioè moti che osservati su opportune partizioni producono successioni con una statistica uguale a quella delle successioni di uno schema di Bernoulli, che è il prototipo della "caoticità e casualità".

La già accennata connessione fra l' entropia e l' instabilità dei moti fa pensare che sistemi dotati di moti instabili tendano ad avere entropia positiva e quindi, per quanto ora detto, moti caotici nel senso più forte possibile del termine. La precisa connessione fra l' entropia e gli esponenti di Lyapunov di un sistema dinamico (M, S, μ) definito su una superficie M è stata recentemente approfondita proprio in connessione all' interesse della questione per la fisica sperimentale. Ma non possiamo qui addentrarci in questa analisi.

æ

VIRIALE (serie del)

per il DIZIONARIO DELLE SCIENZE FISICHE
Autore della voce Giovanni Gallavotti
edito dall'Istituto dell'Enciclopedia Italiana
agosto 1985

Viriale: serie ed equazione di Van der Waals

Accenniamo qui ad una applicazione semplice della meccanica statistica Classica, che pure ne ha segnato un punto di crisi. Una crisi non concettualmente così grave come quelle discusse nella voce Equipartizione, perché solo legata ad una troppo ingenua speranza che la teoria delle transizioni di fase fosse facilmente accessibile anche da un punto di vista quantitativo.

Si considera un gas reale e di N particelle identiche di massa m racchiuse in un contenitore di volume V; si assume, perché così si ottengono buoni risultati empirici, un modello di "Lennard–Jones" per il potenziale microscopico descrivente l' interazione fra due particelle:

$$\varphi(r) = 4\varepsilon((r_0/r)^{12} - (r_0/r)^6) \tag{1}$$

ove ε è l' "intensità" della interazione (con dimensione di energia) e r_0 è il "diametro" delle molecole.

Ci proponiamo di calcolare l' equazione di stato a bassa densità, $V/N \equiv v \longrightarrow \infty$, calcolando la funzione di partizione canonica. Sia:

$$T(\underline{p}) = \sum_{i=1}^{n} \underline{p}_i^2/2m, \qquad \Phi(\underline{q}) = \sum_{i<j} \varphi(\underline{q}_i - \underline{q}_j) \tag{2}$$

e si denoti $\beta = 1/kT$ e $\lambda = \sqrt{2\pi\, m\, k\, T/h^2}^3$, con $k =$ costante di Boltzmann, $T =$ temperatura assoluta, $h =$ costante di Planck; si trova:

$$\begin{aligned} Z(\beta, V) &= \int e^{-\beta(T(\underline{p})+\Phi(\underline{q}))} \frac{d\underline{p}d\underline{q}}{/h^{3N}N!} = \\ &= \frac{1}{N!}\Big(\frac{2\pi mkT}{h^2}\Big)^{3N/2} \int e^{-\beta\Phi(\underline{q})} d\underline{q} = \frac{\lambda^N}{N!} \int \prod_{i<j} e^{-\beta\varphi(\underline{q}_i - \underline{q}_j)} d\underline{q} \end{aligned} \tag{3}$$

si veda la voce "Insiemi Statistici", §2.

In ognuno di fattori del prodotto in (3) si aggiunge e sottrae 1 e poi si sviluppa il prodotto. Posto:

$$f_{ij} = f(r_{ij}) = e^{-\beta\varphi(\underline{q}_i - \underline{q}_j)} - 1, \qquad r_{ij} = |\underline{q}_i - \underline{q}_j|$$

si trova:

$$Z(\beta, V) = \frac{\lambda^N}{N!}\Big(\int \big(1 + \sum_{i<j} f_{ij} + \sum_{\substack{i_1<j_1 \\ (i_1,j_1)}} \sum_{\substack{i_2<j_2 \\ \neq(i_1,j_2)}} f_{i_1 j_1} f_{i_2 j_2} + \ldots\big) dq\Big) =$$

$$= \frac{\lambda^N}{N!}\big(V^N + \binom{N}{2} V^{N-1} \int f(r) 4\pi r^2 dr + \ldots\big) \; . \tag{4}$$

In generale il termine di grado p nelle f diverge con V per $V \longrightarrow \infty$ come $V^N V^p$, (si osservi che in questo calcolo le potenze di N equivalgono a potenze di V, perché $V/N = v =$ costante), ed il suo coefficiente può essere espresso come somma di vari termini ordinati in potenze crescenti di V a partire da V e introducendo:

$$I = (1/2) \int (\exp -\beta\varphi(r) - 1) 4\pi r^2 dr \tag{5}$$

si trova, con un po' di riflessione e scrivendo per ogni ordine in $V \to \infty$, che $Z(\beta, V)$ è:

$$\frac{\lambda^N}{N!} V^N (1 + I\frac{V}{v} + \frac{1}{2}(I\frac{V}{v})^2 + \frac{1}{3!}(I\frac{V}{v})^3 + \ldots) \tag{6}$$

e quindi, usando la formula di Stirling per $N!$:

$$-\beta f_c(\beta, v) = \lim_{V\to\infty} \frac{1}{N} \log \frac{\lambda^N V^N}{N!} e^{V(Iv^{-2} + O(v^{-3}))} =$$

$$= \log v + \log \lambda e + I v^{-1} + O(v^{-2}) \tag{7}$$

Il calcolo di I può essere eseguito approssimativamente, se $\beta\varepsilon \ll 1$ (ossia "ad alta temperatura"), immaginando che $\varphi(r) = +\infty$ (e cioè $f(r) = -1$) per $r < r_0$ e $f(r) = -\beta\varphi(r)$ per $r > r_0$. Si ha:

$$I \cong \frac{1}{2} \int_0^{r_0} -4\pi r^2 dr - \frac{\beta}{2} \int_{r_0}^{\infty} \varphi(r) 4\pi r^2 dr =$$

$$= -4v_0 + \frac{32}{3}\beta\varepsilon v_0 = -(b - \beta a) \tag{8}$$

con:

$$b = 4v_0, \quad a = \frac{32}{3}\varepsilon v_0, \quad v_0 = \frac{4\pi}{3}(\frac{r_0}{2})^3 \tag{9}$$

Allora dalla (7) si trova:

$$\beta P = -\beta \frac{\partial f_c}{\partial v}(\beta, v) = v^{-1} + (b - \beta a)v^{-2} + O(v^{-3}) \tag{10}$$

che fornisce l' equazione di stato a meno di $O(v^{-3})$ e per $\beta\varepsilon \ll 1$, cioè ad alta temperatura. È possibile calcolare, o almeno dare le formule, per i coefficienti di tutti gli ordini della "serie del viriale":

$$\beta P = v^{-1} + \sum_{p=2}^{\infty} c_p(\beta) v^{-p} \tag{11}$$

e anzi si può addirittura dimostrare che la serie converge per v grande, se la funzione φ è data dalla (1) o, più in generale, se verifica le ipotesi di stabilità e temperatezza discusse al §2 della voce Insiemi Statistici, (2.17), (2.18).

La (10) può essere confrontata con una ben nota equazione di stato empirica, la equazione di van der Waals:

$$\beta(P + a/v^2)(v - b) = 1 \quad \text{o} \quad (P + An^2/V^2)(V - nB) = nRT \tag{12}$$

ove, denotando N_A il numero di Avogadro:

$$A = aN_A^2, \quad B = bN_A, \quad R = kN_A, \quad n = N/N_A \ . \tag{13}$$

È chiaro che (11), (12) coincidono a meno di $O(v^{-3})$ e quindi le (13), (9) ci fanno vedere come sia possibile risalire ai parametri microscopici ε e r_0 del potenziale φ (una volta che si è assunta per φ una forma specifica) a mezzo di misure sul gas rarefatto rivelanti le deviazioni dalla legge di Boyle–Mariotte, $\beta P v = 1$:

$$\varepsilon = 3a/8b = 3A/8BN, \quad r_0 = (3b/2\pi)^{1/3} = (3B/2\pi N_A)^{1/3} \tag{14}$$

L' equazione (12) viene però empiricamente usata al di là della sua regione di validità misurando i valori di A e B nell' intorno di valori generici di P, V, T. Il risultato è che i valori di A e B non variano troppo e, accettando questa piccola variabilità di A e B in funzione di v e T, le previsioni della (12) sono state in discreto accordo con l' esperienza, finché al crescere della precisione degli esperimenti non ne sono venute in luce gravi inadeguatezze.

Precisamente la (12) prevede che il gas subisca una transizione di fase di liquefazione con punto critico ad una temperatura T_c e volume V_c e pressione P_c legati da (12) e dati da:

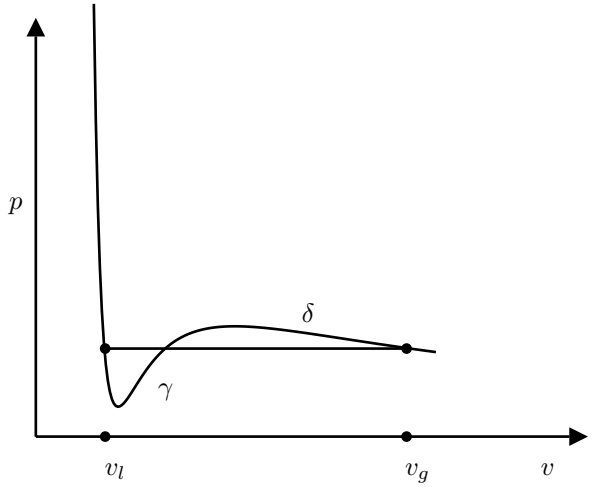
$$RT_c = 8A/27B, \qquad V_c = 3B \qquad (n = 1) \tag{15}$$

Da un lato questo è molto interessante perché fa vedere che esistono semplici relazioni fra i parametri critici e le costanti di interazione microscopiche ($\varepsilon \simeq kT_c$ e $r_0 \simeq (V_c/N_a))^{1/3}$:

$$\varepsilon = 81kT_c/64, \qquad r_0 = (V_c/2\pi N_A)^{1/3} \tag{16}$$

se si usa il modello (1) per il potenziale di interazione φ.

D'altro canto l' equazione (12) non può essere accettabile, almeno per $T < T_c$ perché P non è più funzione monotona di v, come invece dovrebbe essere perché, dalla teoria degli insiemi statistici si può dedurre che l' energia libera (ad esempio calcolata nell' insieme canonico) $f_c(\beta, v)$ è necessariamente convessa in v e cioè $\partial^2 f_c/\partial v^2 \geq 0$ e, quindi, $-\partial P/\partial v = \partial^2 f_c/\partial v^2 \geq 0$.



(17)

Se, tuttavia, le isoterme di (12) vengono prese sul serio anche per $T < T_c$ interpretandole come equazioni di stato per stati metastabili, allora la vera equazione di stato può essere ottenuta dicendo che P in funzione di v ha un tratto orizzontale $[v_l, v_g]$ nelle situazioni in figura, ove un tale tratto associato all' isoterma rappresentata è disegnato; e quindi la densità subisce un salto al decrescere della pressione da v_l a v_g, interpretabili come volumi specifici del liquido e del gas.

Il tratto orizzontale va tracciato in modo tale che le aree γ e δ siano uguali: perché il ciclo ottenuto percorrendo il tratto orizzontale e poi tornando indietro lungo i tratti curvi della stessa isoterma è un ciclo di Carnot a temperatura costante e la differenza delle aree rappresenterebbe il lavoro da esso compiuto (nullo per il secondo principio della termodinamica).

È questa la celebre "costruzione di Maxwell" che, come si vede è motivata in modo piuttosto oscuro poiché non è chiaro se il ciclo di Carnot sia veramente possibile dal momento che è per lo meno dubbio che gli stati "intermedi", ove p è crescente in v, possano essere realmente osservati sperimentalmente o possibili teoricamente, si vedano le voci Transizioni di Fase e Metastabilità.

L' equazione di Van der Waals, perfezionata dalla costruzione di Maxwell, fornisce una semplice rappresentazione della transizione liquido gas. Ma prevede i seguenti comportamenti:

$$\begin{aligned} &(P - P_c) \propto (V - V_c)^{\delta} \qquad \delta = 3,\ T = T_c \\ &(v_g - v_l) \propto (T_c - T)^{\beta} \qquad \beta = 1/2, \text{ per } T \to T_c^- \end{aligned} \tag{17}$$

che sono in contrasto con i dati sperimentali delle più semplici sostanze ($\delta \cong 5$, $\beta \cong 1/3$).

La misura accurata di δ e β è molto delicata e questo spiega perché a lungo l' equazione di Van der Waals fu ritenuta una buona rappresentazione anche della transizione liquido gas.

Per un' idea dei valori di A e B si veda la voce Meccanica Statistica Classica, §2, tavola 1.

Va sottolineato che il mancato accordo con l' esperienza ora menzionato è qui dovuto a cause altre che quelle discusse nella voce Equipartizione e Critica dei Fondamenti: si tratta di un difetto dovuto a cattive approssimazioni (quali l' aver trascurato i termini di ordine superiore in $1/v$ nelle (6), (7) o l' aver supposto che la serie del viriale converga ancora per valori di v, T vicini a quelli del punto critico).

Questo disaccordo non coinvolge nessun problema sui fondamenti della meccanica statistica, si veda la voce Transizioni di fase.

In realtà l' equazione di Van der Waals rappresenta "rigorosamente" solo una situazione limite in cui le particelle hanno cuore duro ed interagiscono con un potenziale φ molto piccolo ma di portata estremamente lunga: si veda la voce Transizioni di fase.

Instabilities and Phase Transitions in the Ising Model. A Review.

Articolo apparso su La Rivista del Nuovo Cimento, **2**, *133–169, 1972; autore: Giovanni Gallavotti*

(1) Introduction.
(2) The model. Grand canonical and canonical ensembles. Their inequivalence.
(3) Boundary conditions. Equilibrium states.
(4) The Ising model in 1 and 2 dimensions and zero field.
(5) Phase transitions. definitions.
(6) Geometric description of the spin configurations.
(7) Phase transitions. Existence.
(8) Microscopic description of pure phases.
(9) Results on phase transitions in a wider range of temperature.
(10) Separation and coexistence of pure phases. Phenomenological considerations.
(11) Separation and coexistence of phases. results.
(12) Surface tension in two dimensions. Alternative descriptions of the separation phenomena.
(13) The structure of the line of separation. What a straight line really is.
(14) Phase separation phenomena and boundary conditions. Further results.
f (15) Conclusions and open problems.
Appendix: Transfer matrix in the Ising model.

Questo articolo è quì riprodotto senza variazioni o aggiornamenti. Sebbene la maggior parte dei problemi elencati come problemi aperti sono stati oggi risolti mi pare che rifletta ancora il punto di vista corrente sulla teoria delle transizioni di fase e dei fenomeni di coesistenza e fornisca una corretta analisi dei fondamenti di queste teorie. Ho aggiunto alcune note bibliografiche che mi sono parse irrinunciabili (nella forma di note a piè pagina), senza pretesa di completezza.

Roma, march 1972

§1. – Introduction.

The Ising model plays a very special role in statistical mechanics and provides the simplest non trivial example of a system undergoing phase transitions [1].

The analysis of this model has provided deep insight into the general nature of the phase transitions which are certainly better understood nowadays after the publication of the hundreds of papers which followed the pioneering work of Ising, Peierls, Onsager, and Lee and Yang, [1-4].

The main reason why so much attention has been given to this very special model lies in its simplicity and in the fact that, in spite of it, it first provided firm and quantitative indications that a microscopic short-ranged interaction can produce phase transitions which, furthermore, deeply differ in character from the classical Van der Waals (or Curie-Weiss or mean field) type of transitions [5].

It should also be mentioned that the two-dimensional Ising model in zero external field is exactly solvable [6], this fact has very often been used as a check of the validity of numerical approximations devised to be applied to more complicated models [7].

In some cases the Ising model is a good phenomenological model for antiferromagnetic materials [8].

Finally, last but not least, we mention that the Ising model has given rise to a number of interesting developments and reinterpretations of old results in the theory of Markov chains [9], information theory [10], random walks [11], and therefore constitutes a remarkable example of a subject which has simoultaneously been the object of advanced research in physics, mathematics and mathematical physics.

In this review we hope to give a self contained, though certainly not exhaustive, description of the model and of some selected rigorous results. They illustrate properties which throw some light on the general nature of the phenomenon of phase transition far from the critical point and which, hopefully, should not be a peculiarity of the simplicity of the model.

There exist some very good accounts on the theoretical arguments leading to the consideration of the Ising model in the context of physical problems [7,12]. Therefore we shall completely skip this aspect of the matter and refer the interested reader to the literature.

§2. – The model. Grand canonical and canonical ensembles. Their inequivalence.

We consider a δ-dimensional ($\delta = 1, 2, 3$) square lattice $\mathbf{Z}^\delta$ and a finite square $\Lambda \subset \mathbf{Z}^\delta$ centred around the origin, containing $|\Lambda| = L^\delta$ lattice sites.

On each site $x \in \Lambda$ is located a classical spin $\sigma_x = \pm 1$. The "configurations" of our system will, therefore, consist in a set $\underline{\sigma} = (\sigma_{x_1}, \ldots, \sigma_{x_{|\Lambda|}})$ of $|\Lambda|$ numbers $\sigma_x = \pm 1$; the number of these configurations is $2^{|\Lambda|}$. The ensemble of the configurations will be denoted $\mathcal{U}(\Lambda)$.

To each spin configuration is assigned a certain "energy":

$$H_\Lambda(\underline{\sigma}) = -J \sum_{<i,j>} \sigma_{x_i}\sigma_{x_j} - h \sum_i \sigma_{x_i} - \mathcal{B}_\Lambda(\underline{\sigma}) \tag{2.1}$$

where $\sum_{<i,j>}$ means that the sum is over pairs (x_i, x_j) of neighbouring points, h is an external magnetic field and $\mathcal{B}_\Lambda(\underline{\sigma})$ describes the interaction of the spins in the box Λ with the "rest of the world" [13].

For simplicity we shall treat, in this paper, only the case $J > 0$.

Of course $\mathcal{B}_\Lambda(\underline{\sigma})$ in (2.1) can be rather arbitrary and, actually, depends on the particular physical problem under investigation. It is subject, however, to one constraint of physical nature: in case we were interested in letting $\Lambda \to \infty$, we should impose the condition:

$$\lim_{\Lambda\to\infty} \frac{\max_{\underline{\sigma}} |\mathcal{B}_\Lambda(\underline{\sigma})|}{|\Lambda|} = 0 \tag{2.2}$$

i.e. we want that the energy due to $\mathcal{B}_\Lambda(\underline{\sigma})$ should not be of the same order as the volume of the box. In other words $\mathcal{B}_\Lambda$ should be a "surface term".

The laws of statistical mechanics provide a relationship between the microscopic Hamiltonian (2.1) and the macroscopic quantities appearing in the thermodynamical theory of the system.

The free energy per unit volume is given by:

$$f_\Lambda(\beta, h) = \frac{\beta^{-1}}{|\Lambda|} \log Z(\beta, h, \Lambda, \mathcal{B}) \tag{2.3}$$

where $\beta = T^{-1}$ is the inverse temperature and:

$$Z(\beta, h, \Lambda, \mathcal{B}) = \sum_{\underline{\sigma}\in\mathcal{U}(\Lambda)} e^{-\beta H_\Lambda(\underline{\sigma})} \tag{2.4}$$

is the grand canonical partition function. Furthermore the probability of finding the system in a configuration $\underline{\sigma}$ of the grand canonical ensemble $\mathcal{U}(\Lambda)$ is given by the Boltzmann factor:

$$\frac{e^{-\beta H_\Lambda(\underline{\sigma})}}{Z(\beta, h, \Lambda, \mathcal{B})}, \qquad \underline{\sigma} \in \mathcal{U}(\Lambda) \tag{2.5}$$

For a theoretical foundation of (2.3), (2.5) see [14]. The grand–canonical ensemble formalism based on (2.3),(2.5) corresponds to the physical situation in which there are no constraints on the system. If one could, by

some experimental arrangement, regard for example the total magnetization $M(\underline{\sigma}) = \sum_{x\in\Lambda}\sigma_x$ as fixed: $M(\underline{\sigma}) = M = m|\Lambda|$, then the expression (2.3) for the free energy would no longer be appropriate.

One should rather consider the canonical ensemble, *i.e.* the set of the allowed configurations would be the set $\mathcal{U}(\Lambda, m) \subset \mathcal{U}(\Lambda)$ consisting of all the $\underline{\sigma} \in \mathcal{U}(\Lambda)$ such that $\sum_{x\in\Lambda}\sigma_x = m|\Lambda|$, ($|m| < 1$), and the thermodynamics would be described by the function:

$$g_\Lambda(\beta, h, m) = \frac{\beta^{-1}}{|\Lambda|}\log Z(\beta, h, \Lambda, \mathcal{B}, m) \tag{2.6}$$

where:

$$Z(\beta, h, \Lambda, \mathcal{B}, m) = \sum_{\underline{\sigma}\in\mathcal{B}(\Lambda,m)} e^{-\beta H_\Lambda(\underline{\sigma})} \tag{2.7}$$

and the free energy would be $\tilde{f}_\Lambda(\beta, h)$:

$$\tilde{f}_\Lambda(\beta, h) = h\, m(h) + g_\Lambda(\beta, 0, m(h)) \tag{2.8}$$

where $m(h)$ is the solution of the equation [15]:

$$h = -\frac{\partial g_\Lambda(\beta, 0, m)}{\partial m} \tag{2.9}$$

There is no reason for having $\tilde{f}_\Lambda = f_\Lambda$ since they correspond to different physical problems; it is only when, *in some sense*, the fluctuations become neglegible (*i.e.* in the limit $\Lambda \to \infty$) that one can expect the identity between $\tilde{f}$ and f.

Of course in general the difference between $\tilde{f}_\Lambda$ and f_Λ should vanish as $|\Lambda|^{-1}$ times $O(|\Lambda|^{(\delta-1)/\delta})$ ($O(\log|\Lambda|)$ for $\delta = 1$); but, as we shall see on many occasions, the situtation is not so simple for other quantities such as the correlation functions or the average magnetization.

The inequivalence, for finite volume, of the predictions of the canonical and grand canonical ensembles should not be interpreted as meaning that statistical mechanics is only approximate when applied to finite systems; it simply means that in dealing with finite systems care must be paid not only to the boundary conditions but also to the actual physical situation from which the problem under consideration arises.

We conclude by remarking that in the grand canonical ensemble the probability of a spin configuration will be given by an expression similar to (2.5):

$$\frac{e^{-\beta H_\Lambda(\underline{\sigma})}}{Z(\beta, h, \Lambda, \mathcal{B}, m)}, \qquad \underline{\sigma} \in \mathcal{U}(\Lambda, m) \tag{2.10}$$

§3. – Boundary conditions. Equilibrium states.

Formula (2.5), or (2.10), provide a complete statistical description of the properties of the system. An alternative and often more convenient, equally complete, description is provided by the so-called correlation functions:

$$\langle \sigma_{x_1}\sigma_{x_2}\dots\sigma_{x_n}\rangle_{\Lambda,\mathcal{B}_\Lambda} = \frac{\sum_{\underline{\sigma}} \sigma_{x_1}\sigma_{x_2}\dots\sigma_{x_n}\, e^{-\beta H_\Lambda(\underline{\sigma})}}{\sum_{\underline{\sigma}} e^{-\beta H_\Lambda(\underline{\sigma})}} \tag{3.1}$$

where the $\sum_{\underline{\sigma}}$ is extended to the appropriate statistical ensemble.

For instance the average magnetization in the grand canonical ensemble $\mathcal{U}(\Lambda)$ is:

$$m_\Lambda(\beta,h) = \frac{\partial f_\Lambda(\beta,h)}{\partial h} = \frac{\sum_{x\in\Lambda}\langle\sigma_x\rangle_{\Lambda,\mathcal{B}}}{|\Lambda|} \tag{3.2}$$

We shall refer to the family of correlation functions (3.1) (regarded as a whole) as the "equilibrium state of the system in the box Λ".

We call equilibrium state of the infinite system any family $\{\langle\sigma_{x_1}\dots\sigma_{x_n}\rangle\}$ of functions such that, for a suitable choice of the $\mathcal{B}_\Lambda(\underline{\sigma})$:

$$\langle\sigma_{x_1}\dots\sigma_{x_n}\rangle = \lim_{\Lambda\to\infty}\langle\sigma_{x_1}\dots\sigma_{x_n}\rangle_{\Lambda,\mathcal{B}_\Lambda} \tag{3.3}$$

for all $n \geq 1$ and all $x_1, x_2, \dots, x_n \in \mathbf{Z}^\delta$, *simoultaneously* [16].

An equilibrium state for an infinite system will simply be called an equilibrium state: it is specified by a suitable choice of a sequence $\{\mathcal{B}_\Lambda(\underline{\sigma})\}$ of boundary conditions satisfying the requirement (2.2).

Let us list a number of remarkable boundary conditions:

1) *Open boundary condition* (also called "perfect–wall" boundary conditions): this name will be given to the case:

$$\mathcal{B}_\Lambda(\underline{\sigma}) \equiv 0 \qquad \text{for all } \underline{\sigma} \in \mathcal{U}(\Lambda) \tag{3.4}$$

2) *Periodic boundary conditions*: this corresponds to allowing spins on opposite faces of the box Λ to interact through a coupling $-J$ (*i.e.* as the the bulk spins). Clearlt this can be obtained by a suitable choice of $\mathcal{B}_\Lambda(\underline{\sigma})$; to this choice we hall refer as "periodic boundary conditions".

3) $(\underline{\varepsilon})$-*boundary conditions*: let $(\xi_1,\xi_2,\dots)$ be the $2\delta|\Lambda|^{(\delta-1)/\delta}$ lattice points adjacent to the boundary of Λ. Let $\underline{\varepsilon} = (\varepsilon_{\xi_1},\varepsilon_{\xi_2},\dots)$, $\varepsilon_{\xi_i} = \pm 1$, be fixed. We shall call (ε)-boundary condition the choice:

$$\mathcal{B}_\Lambda(\underline{\sigma}) = -J\sum_{x_i\in\partial\Lambda}\sigma_{x_j}\varepsilon_{\xi_j} \tag{3.5}$$

where (x_i, ξ_i) are nearest neighbours.

The physical meaning of this boundary condition is clear: we imagine that the sites neighbouring the boundary $\partial\Lambda$ of Λ are occupied by a spin configuration $\underline{\varepsilon}$ and that the latter spins interact with the spins $\underline{\sigma}$ through the same coupling constant of the bulk spins.

The cases $\underline{\varepsilon} = (+1, +1, \ldots, +1)$ or $\underline{\varepsilon} = (-1, -1, \ldots, -1)$ will be, respectively, referred to as the $(+)$-*boundary condition* or the $(-)$-*boundary condition*.

4) In the two-dimensional case we shall be interested in another boundary condition. Suppose the spins on the opposite *vertical* sides of Λ are allowed to interact through a coupling $-J$ (*i.e.* we impose periodic boundary conditions along the rows of Λ *only*); and suppose that a set $\underline{\varepsilon}_u$ of fixed spins is located on the lattice sites adjacent to the upper base of Λ and, similarly, a set $\underline{\varepsilon}_l$ of fixed spins is adjacent to the lower base of Λ. The spins $\underline{\varepsilon}_u, \underline{\varepsilon}_l$ are allowed to interact with the nearest spins in Λ with a coupling $-J$. We shall naturally refer to this choice of $\mathcal{B}_\Lambda(\underline{\sigma})$ as the $(\underline{\varepsilon}_u, \underline{\varepsilon}_l)$-cylindrical boundary condition.

The particular cases:

$$\underline{\varepsilon}_u = (+1, +1, \ldots, +1), \qquad \underline{\varepsilon}_l = (+1, +1, \ldots, +1)$$

or

$$\underline{\varepsilon}_u = (+1, +1, \ldots, +1), \qquad \underline{\varepsilon}_l = (-1, -1, \ldots, -1)$$

will be referred to respectively as $(+,+)$-*cylindrical boundary* condition or $(+,-)$-*cylindrical boundary condition*.

§4. – The Ising model in 1 and 2 dimensions and zero field.

To acquire some familiarity with the model let us examine some of the simplest cases.

Consider the one-dimensional Ising chain with periodic boundary conditions. If we label the points of Λ as $1, 2, \ldots, L$, the Hamiltonian in zero field is:

$$H_\Lambda(\underline{\sigma}) = -J \sum_{i=1}^{L} \sigma_i \sigma_{i+1}, \qquad \sigma_{L+1} \equiv \sigma_1 \tag{4.1}$$

(clearly $\mathcal{B}_\Lambda(\underline{\sigma}) = -J\sigma_L\sigma_1$). The gran canonical partition function can be written:

$$Z_\Lambda(\beta) = \sum_{\underline{\sigma}} e^{\beta J \sum_{i=1}^{L} \sigma_i \sigma_{i+1}} = \sum_{\underline{\sigma}} \prod_{i=1}^{L} e^{\beta J \sigma_i \sigma_{i+1}} \tag{4.2}$$

Remarking that $(\sigma_i\sigma_{i+1})^2 \equiv 1$ and, therefore:

$$e^{\beta J\sigma_i\sigma_{i+1}} \equiv \cosh\beta J - \sigma_i\sigma_{i+1}\sinh\beta J$$

eq. (4.2) can be rewritten as:

$$Z_\Lambda(\beta) = (\cosh\beta J)^L \sum_{\underline{\sigma}} \prod_{i=1}^{L} (1 + \tanh\beta J\,\sigma_i\sigma_{i+1}) \tag{4.3}$$

If one develops the product in (4.3) one gets a sum of terms of the form:

$$(\tanh\beta J)^k \sigma_{i_1}\sigma_{i_1+1}\sigma_{i_2}\sigma_{i_2+1}\ldots\sigma_{i_k}\sigma_{i_k+1} \tag{4.4}$$

It is clear that, unless $k = 0$ or $k = L$, each of the terms (4.4) contains at least an index i_j which appears only once. Therefore, after performing the sum over the $\underline{\sigma}$'s, all the terms (4.4) give a vanishing contribution to $Z_\Lambda(\beta)$ except the two with $k = 0$ and $k = L$ which are, respectively, 1 and $(\tanh\beta J)^L \cdot \sigma_1\sigma_2\sigma_2\ldots\sigma_{L-1}\sigma_L\sigma_L\sigma_1 \equiv (\tanh\beta J)^L$.
This implies:

$$Z_\Lambda(\beta) = (\cosh\beta L)^L 2^L\,(1 + (\tanh\beta J)^L) \tag{4.5}$$

Hence [17]:

$$\beta f_\Lambda(\beta) = \log(2\cosh\beta J) + \frac{1}{L}\log(1 + (\tanh\beta J)^L) \tag{4.6}$$

It has to be remarked that $\beta f_\Lambda(\beta)$ as well as $\beta f(\beta) = \lim_{L\to\infty}\beta f_\Lambda(\beta) = \log 2\cosh\beta J$ is analytic in β; this fact is usually referred to as the absence of phase transitions in the one-dimensional Ising model.
The reader can check, by using the above scheme, that the partition function in the grand canonical ensemble and zero field but open boundary conditions (see Sect. **3**) is slightly different from (4.5) and, precisely, is equal to $(\cosh\beta J)^{L-1}2^L$.
Consider now the two-dimensional Ising model in a zero field and with open boundary conditions:

$$H_\Lambda(\underline{\sigma}) = -J\sum_{i=1}^{L}\sum_{j=1}^{L-1}\sigma_{i\,j}\sigma_{i\,j+1} - J\sum_{i=1}^{L-1}\sum_{j=1}^{L}\sigma_{i\,j}\sigma_{i+1\,j} \tag{4.7}$$

A better form for $H_\Lambda(\underline{\sigma})$ is the following:

$$H_\Lambda\underline{\sigma} = -J\sum_{b}\tilde{\sigma}_b \tag{4.8}$$

where $\sum_b$ denotes sum over the bonds, *i.e.* over the segments $b = [(i,j),(i,j+1)]$ or $b = [(i,j),(i+1,j)]$, and $\tilde{\sigma}_b$ is the product of the two spins at the extremes of b (*e.g.* , if $b = [(i,j),(i+1,j)]$ then $\tilde{\sigma}_b = \sigma_{i\,j}\sigma_{i+1,j}$).

The partition function can be written, as in the one-dimensional case, as:

$$Z_\Lambda(\beta) = (\cosh \beta J)^{2L(L-1)} \sum_{\underline{\sigma}} \prod_b (1 + (\tanh \beta J)\tilde{\sigma}_b) \tag{4.9}$$

Developing the product we are led to a sum of terms of the type:

$$(\tanh \beta J)^k \tilde{\sigma}_{b_1} \tilde{\sigma}_{b_2} \dots \tilde{\sigma}_{b_k} \tag{4.10}$$

and we can conveniently describe this term through the geometric set of lines $b_1, b_2, \dots, b_k$. After the $\sum_{\underline{\sigma}}$ is taken, many terms of the form (4.1) give a vanishing contribution. The ones that give a non vanishing contribution are the ones in which the vertices of the geometric figure $b_1 \cup b_2 \cup \dots \cup b_k$ belong to an even number of b_j's (two or four). These terms are the ones such that $\tilde{\sigma}_{b_1} \cdot \tilde{\sigma}_{b_2} \dots \tilde{\sigma}_{b_k} \equiv 1$. In Fig. 1 we give a typical non vanishing term and in Fig. 2 an example of a vanishing term ($k = 30$).

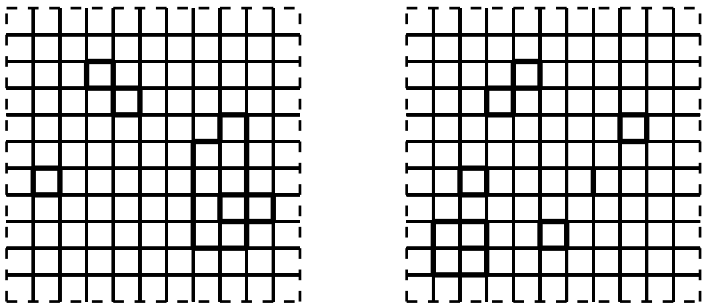


Fig. 1,2 - The dashed line is the boundary of Λ.

We shall, in the following, consider the geometric figures built with k segments $b_1, \dots, b_k$ such that $\tilde{\sigma}_{b_1} \cdot \tilde{\sigma}_{b_2} \dots \tilde{\sigma}_{b_k} \equiv 1$ and call it a *k-sided multipolygon* on the box Λ (needless to say that all the $b_1, \dots, b_k$ are different from each other). Let $P_k(\Lambda)$ be the number of such polygons.

The partition function is now easily written as [18]:

$$Z_\Lambda(\beta) = (\cosh \beta J)^{2L(L-1)} 2^{L^2} \sum_{k \geq 0} P_k(\Lambda) \, (\tanh \beta J)^k \tag{4.11}$$

§5 – Phase transitions. Definitions.

We have already seen, in the preceding Section, that the one dimensional Ising model has no phase transitions in zero field since both $f_\Lambda(\beta)$ and $f(\beta)$ are analytic in β.

We wish to discuss in more detail what is meant by a "phase transition". It should be said at the ouset that there is no universally agreed upon definition of such a concept. Intuitively, from everyday experience, one would say that a phase transition is a phenomenon of macroscopic instability: slight changes of external conditions should imply dramatic changes of some macroscopic variables; it is hard to imagine how in such a situation thermodynamic functions like the free energy, *etc*, could be analytic functions of the parameters in terms of which they are expressed (say, temperature, chemical potential or magnetic field, *etc*).

For the above reason an analytic singularity in the thermodynamic functions is usually thought as a "symptom" of a phase transition and on this idea it would be possible to base a definition and a theory of the phenomenon of phase transitions.

In this paper, however, we will not base the investigation of the nature of phase transitions in the Ising model on the search of singularities of the thermodynamic functions; we shall rather adopt and make more precise the other, perhaps more immediate and intuitive, approach based on the detection of "macroscopic instabilities".

This way of proceeding is more convenient for the simple reason that a number of very clear and rather deep results have beem obtained along these lines. But it should be understood that this second approach does not "brilliantly" avoid the difficulties of the first. It is simply an approach to the theory of phase transitions which, so far, has asked and provided a less refined description of the phenomena of interest, as compared to the description which would be expected from the analysis of the singularities of appropriate analytic functions (an analysis still in a very primitive stage and whose problems are often not well formulated even in the simplest cases) [19].

Let us now discuss in a more precise way the concept of macroscopic instability.

Consider the Ising model and define that a phase transition takes place at the values (β, h) of the thermodynamic parameters if the system is unstable with respect to boundary perturbations; *i.e.* if there are at least two sequences $\mathcal{B}_\Lambda(\underline{\sigma})$ and $\mathcal{B}'_\Lambda(\underline{\sigma})$ of boundary terms (see (2.1), (2.2)) such that (say, in the grand canonical ensemble):

$$\lim_{\Lambda\to\infty} \langle \sigma_{x_1} \dots \sigma_{x_n} \rangle_{\Lambda,\mathcal{B}_\Lambda} \neq \lim_{\Lambda\to\infty} \langle \sigma_{x_1} \dots \sigma_{x_n} \rangle_{\Lambda,\mathcal{B}'_\Lambda} \tag{5.1}$$

for a suitable choice of $x_1, x_2, \dots, x_n, n$.

We first clarify why we say that, if (5.1) holds, we have a macroscopic instability.

We remark that a change in the boundary conditions does not change the extensive properties of the system such as the free energy. In fact, from the definition (2.4):

$$\frac{Z(\beta,h,\Lambda,\mathcal{B}_\Lambda}{Z(\beta,h,\Lambda,\mathcal{B}'_\Lambda)} \leq e^{\max_{\underline{\sigma}\in\mathcal{U}(\Lambda)} |\mathcal{B}_\Lambda(\underline{\sigma})|+|\mathcal{B}'_\Lambda(\underline{\sigma})|} \tag{5.2}$$

and therefore (2.2) implies:

$$\lim_{\Lambda\to\infty} \frac{1}{|\Lambda|} \log Z(\beta,h,\Lambda,\mathcal{B}_\Lambda) \equiv \lim_{\Lambda\to\infty} \frac{1}{|\Lambda|} Z(\beta,h,\Lambda,\mathcal{B}'_\Lambda) \tag{5.3}$$

On the other hand, if (5.1) is true, intensive quantities like the correlation functions are sensitive to the boundary conditions; for instance if:

$$\lim_{\Lambda\to\infty} \langle \sigma_{x_1} \rangle_{\Lambda,\mathcal{B}_\Lambda} \neq \lim_{\Lambda\to\infty} \langle \sigma_{x_1} \rangle_{\Lambda,\mathcal{B}'_\Lambda}$$

we realize that the local magnetization changes as a consequence of a change in the boundary condition *even* if the boundary is very remote.

Of course once provided with a "definition" of what a phase transition is, one has not gone very far. The real question is whether the definition reflects what is physically expected; this implies, in particular, that one should at least be able to prove the existence of a phase transition in the above sense in cases in which one expects a transition. Hopefully the definition and its physical interpretation should alllow one to do more: for instance to provide the tools for a closer description of typical phenomena (like the phase separation).

We end here the above, somewhat philosophical but necessary, discussion and in the coming Sections. We shall describe in some more concrete examples the results that have been obtained in the last decade when the above point of view was starting to be developed, quite independently, by several people.

§6. – Geometric description of the spin configurations.

In this section we introduce a new description of the spin configurations which we shall use to derive in a very elegant way the exact value of the critical temperature in the two-dimensional Ising model. In the following

Sections the geometric representastion, introuced below, will be widely used for other purposes [20].

Consider an Ising model with boundary conditions of the type (3.5) (($\underline{\varepsilon}$)–boundary conditions) or with periodic boundary conditions (see Sect. **2**).

Given a configuration $\underline{\sigma} \in \mathcal{U}(\Lambda)$ we draw a unit segment perpendicular to the centre of each bond b having opposite spins at its extremes (in three dimensions we draw a unit square surface element perpendicular to b). A two dimensional example of this construction is provided by Fig. 3 (($\underline{\varepsilon}$)–boundary condition).

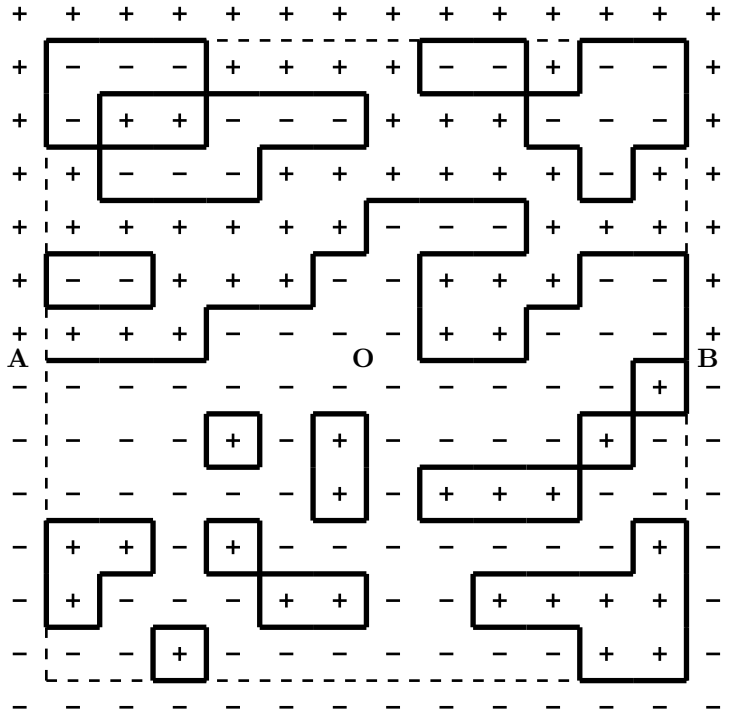


Fig. 3. - The dashed line is the boundary of Λ; the outer spins are the ones fixed by the boundary condition. The points A, B are points where an open line ends.

The set of segments can be grouped into lines (or surfaces in three dimensions) which separate regions where the spins are positive from regions where they are negative.

It is clear that some of the lines (or surfaces, if $\delta = 3$) are "closed polygons" ("closed polyhedra", respectively) while others are not closed. It is perhaps worth stressing that our polygons are not really such in a geometrical sense, since they are not necessarily "self–avoiding" (see Fig. 3): however they are such that they can intersect themselves only on vertices (and not on

sides). From a geometrical point of view a family of disjoint polygons (in the above sense and in two dimensions) is the same thing as a mutipolygon in the sense discussed in Sect. **4** (see Fig. 1).

In two dimensions instead of saying that a polygon is "closed" we could equivalently say that its vertices belong to either two or four sides.

We note that the $(+)$–bundary conditions, the $(-)$–boundary conditions and the periodic boundary conditions are such that the lines (surfaces) associated with spin configurations are *all* closed polygons (polyhedra). In the periodic case some polygons might wind up around the two holes of the torus.

In the two–dimensional case and if the boundary conditions are the $(+,+)$–cylindrical or the $(+,-)$–cylindrical ones (see Sect. **2**) a geometric construction of the above type can still be performed and, also in this case, the lines are closed polygons (some of which may "wind around" the cylinder Λ).

For a fixed boundary condition let $(\gamma_1, \gamma_2, \ldots, \gamma_h, \lambda_1, \ldots, \lambda_k)$ be the disjoint components of the set of lines (surfaces) associated by the above construction with a spin configuration $\underline{\sigma} \in \mathcal{U}(\Lambda)$. The $\gamma_1, \ldots, \gamma_h$ are closed polygons and the $\lambda_1, \ldots, \lambda_k$ are not closed.

Clearly the correspondence between $(\gamma_1, \gamma_2, \ldots, \gamma_h, \lambda_1, \ldots, \lambda_k)$ and $\underline{\sigma}$ is, *for a fixed boundary condition*, one-to-one except for the case of the periodic or open boundary conditions, when it is one–to–two. Changing boundary conditions implies changing the set of lines (surfaces) which describe the same spin configuration $\underline{\sigma}$.

A very important property of the above geometric description is that, if $|\gamma|$, $|\lambda|$ denote the length (area) of the lines (surfaces) γ and λ, then the energy of a spin configuration is, in zero field, given by:

$$H_\Lambda(\underline{\sigma}) = -J \cdot (\text{number of bonds in } \Lambda) + 2J[\sum_i |\gamma_i| + \sum_j |\lambda_j|] \qquad (6.1)$$

This remark easily follows from the fact that each bond b contributing $-J$ to the energy has equal spins at its extremes, while the bonds contributing $+J$ have opposite spins at their extremes and, therefore, are cut by a segment of unit length belonging to some γ_i or λ_j.

If $\mathcal{N}_\Lambda =$ (number of bonds in Λ), the partition function becomes (in zero field and with non periodic boundary conditions):

$$Z_\Lambda(\beta) = \sum_{\gamma_1 \ldots \gamma_h} \sum_{\lambda_1 \ldots \lambda_k} \left(e^{-2\beta J \sum_i |\gamma_i|} \cdot e^{-2\beta J \sum_j |\lambda_j|} \right) \cdot e^{\beta J \mathcal{N}_\Lambda} \qquad (6.2)$$

where the sum runs over the set of lines associated with a spin configuration $\underline{\sigma} \in \mathcal{U}(\Lambda)$ and with the boundary condition under consideration.

In the case of periodic or open boundary conditions there are no λ's and there is an extra factor 2 (due to the two–to–one correspondence between $\underline{\sigma}$ and $(\gamma_1, \ldots, \gamma_h)$):

$$Z_\Lambda(\beta) = 2 \sum_{\gamma_1 \ldots \gamma_h} e^{-2\beta J \sum_i |\gamma_i|} \cdot e^{\beta J \mathcal{N}_\Lambda} \tag{6.3}$$

and $\mathcal{N}_\Lambda = 2L^2$.

Form the above considerations we draw two important consequences:

I) If the boundary condition is fixed, the probability of a spin configuration $\underline{\sigma}$ described by $\gamma_1, \ldots, \gamma_h, \lambda_1, \ldots, \lambda_k$ is proportional to:

$$e^{-2\beta J \left(\sum_i |\gamma_i| + \sum_j |\lambda_j| \right)} \tag{6.4}$$

II) In the case of $(+)$ or $(-)$ boundary conditions and 2 dimensions we remark that $\sum_{\gamma_1 \ldots \gamma_k}$ in (6.2) is a sum over "multipolygons" lying on a shifted lattice and in a box Λ' containing $(L-1)^2$ spins (see definition on Sect. **4** and, therefore, if $\sum_i |\gamma_i| = k$ we have:

$$Z_\Lambda(\beta) = e^{[2L(L-1)\beta J]} \sum_{k \geq 0} P_k(\Lambda')\, e^{-2\beta J k} \tag{6.5}$$

where $P_k(\Lambda')$ is the number of different multipolygons with perimeter k (see (4.11)).

If we now define β^* through:

$$\tanh \beta J = e^{-2\beta^* J} \tag{6.6}$$

then comparison between (6.5) and (4.11) yields:

$$\frac{Z_\Lambda(\beta)}{2^{L^2} (\cosh \beta J)^{2L(L-1)}} = \frac{Z_{\Lambda'}(\beta^*)}{e^{2\beta^* J L(L-1)}} \tag{6.7}$$

Here $Z_\Lambda(\beta)$ is computed with open boundary conditions, while $Z_{\Lambda'}(\beta^*)$ is computed with $(+)$–boundary conditions.

If we assume that the bulk free energy $f(\beta) = \lim_{\Lambda \to \infty} \frac{1}{|\Lambda|} \log Z_\Lambda(\beta)$ has one and only one singularity as a function of β, for β real, then (6.7) can be used to locate the singularity. In fact it implies:

$$f(\beta) - \log 2 (\cosh \beta J)^2 = -2\beta^* J + f(\beta^*) \tag{6.8}$$

hence a singularity in β can take place only when $\beta = \beta^*$, *i.e.* for $\beta = \beta_{c,O}$ such that:

$$\tanh \beta_{c,O} = e^{-2\beta_{c,O} J} \tag{6.9}$$

which, indeed, has been shown by Onsager [3] to be the exact value of the critical temperature defined as the value of β where $f(\beta)$ is singular (in the sense that its derivative diverges) [20].

In the next Section we outline the theory of the phase transitions in the Ising model as a macroscopic instability and a spontaneous breakdown of the up–down symmetry. We shall concentrate, for geometric reasons, on the two–dimensional Ising model but, unless explicitly stated, the results hold in any dimension $\delta \geq 2$.

§7. – Phase transitions. Existence.

In this Section we shall show tthat the (+)–boundary conditions and the (−)–boundary conditions (see Sect. **3**) produce, if the temperature is low enough, different equilibrium states (see Sect. **3**), *i.e.* for large β the correlation functions are different and the difference does not vanish in the limit $\Lambda \to \infty$ (see (5.1)).

More precisely we shall prove that. if $h = 0$ and β is large enough:

$$\lim_{\Lambda\to\infty} \langle \sigma_x \rangle_{\Lambda,\pm} = \pm m^*(\beta) \neq 0 \tag{7.1}$$

where the index $\pm$ refers to the boundary conditions.

Clearly (7.1) shows that the magnetization is unstable (in zero field and at low temperature) with respect to boundary perturbations. We also remark that by using periodic boundary conditions one would obtain still another result:

$$\lim_{\Lambda\to\infty} \langle \sigma_x \rangle_{\Lambda,\, periodic} = 0, \qquad \text{if } h = 0 \tag{7.2}$$

because $\langle \sigma_x \rangle_{\Lambda,\, periodic} \equiv 0$, if $h = 0$, for symmetry reasons.

After a description of the very simple and instructive proof of (7.1) we shall go further and discuss more deeply the character of the phase transition.

As already remarked, the spin configurations $\underline{\sigma} \in \mathcal{U}(\Lambda)$ are described in terms of closed polygons $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ if the boundary condition is (+) or (−) and the probability of a configuration $\underline{\sigma}$ described by $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ is proportional to (see (6.4)):

$$e^{-2\beta J \sum_i |\gamma_i|} \tag{7.3}$$

Below we identify $\underline{\sigma}$ with $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ (with the boundary condition fixed).

Let us estimate $\langle \sigma_x \rangle_{\Lambda,+}$. Clearly $\langle \sigma_x \rangle_{\Lambda,+} = 1 - 2\, P_{\Lambda,+}(-)$, where $P_{\Lambda,+}(-)$ is the probability that in the site x the spin is -1.

We remark that if the site x is occupied by a negative spin the point x is *inside* some contour γ associated with the spin configuration $\underline{\sigma}$ under consideration. Hence if $\rho(\gamma)$ is the probability that a given contour belongs to the set of contours describing a configuration $\underline{\sigma}$, we deduce:

$$P_{\Lambda,+}(-) \leq \sum_{\gamma o x} \rho(\gamma) \tag{7.4}$$

where $\gamma o x$ means that γ surrounds x.

Let us now estimate $\rho(\gamma)$: if $\Gamma = (\gamma_1, \ldots, \gamma_n)$ is a spin configuration and if the symbol $\Gamma \, comp \, \gamma$ means that the contour Γ is disjoint from $\gamma_1, \ldots, \gamma_n$ (*i.e.* $\{\gamma \cup \Gamma\}$ is a new spin configuration), then:

$$\rho(\gamma) = \frac{\sum_{\Gamma \ni \gamma} e^{-2\beta J \sum_{\gamma' \in \Gamma} |\gamma'|}}{\sum_{\Gamma} e^{-2\beta J \sum_{\gamma' \in \Gamma} |\gamma'|}} \equiv e^{-2\beta J |\gamma|} \frac{\sum_{\Gamma \, comp \, \gamma} e^{-2\beta J \sum_{\gamma' \in \Gamma} |\gamma'|}}{\sum_{\Gamma} e^{-2\beta J \sum_{\gamma' \in \Gamma} |\gamma'|}} \tag{7.5}$$

Before continuing the analysis let us remark that if $\underline{\sigma} = (\gamma, \gamma_1, \gamma_2, \ldots, \gamma_n)$ then $\underline{\sigma}' = (\gamma_1, \gamma_2, \ldots, \gamma_n)$ is obtained from $\underline{\sigma}$ by reversing the sign of the spins inside γ; this can be used to build an intuitive picture of the second equation in (7.5).

Clearly the last ratio in (7.5) does not exceed 1; hence:

$$\rho(\gamma) \leq e^{-2\beta J |\gamma|} \tag{7.6}$$

Calling $p = |\gamma|$ and observing that there are at most 3^p different shapes of γ with perimeter p and at most p^2 congruent γ's containing (in their interior) x, we deduce from (7.4), (7.6):

$$P_{\Lambda,+}(-) \leq \sum_{p=4}^{\infty} p^2 3^p e^{-2\beta J p} \tag{7.7}$$

Hence if $\beta \to \infty$ (*i.e.* the temperature $T \to 0$) this probability can be made as small as we like and, therefore, $\langle \sigma_x \rangle_{\Lambda,+}$ is as close to 1 as we like provided β is large enough. It is of fundamental importance that the closeness of $\langle \sigma_x \rangle_{\Lambda,+}$ to 1 is both x and Λ independent.

A similar argument for the $(-)$–boundary condition, or the remark that $\langle \sigma_x \rangle_{\Lambda,-} = -\langle \sigma_x \rangle_{\Lambda,+}$, allows us to conclude that, at large β, $\langle \sigma_x \rangle_{\Lambda,-} \neq \langle \sigma_x \rangle_{\Lambda,+}$ and the difference between the two quantities is uniform in Λ.

Hence we have completed the proof ("Peierls argument") of the fact that there is a strong instability with respect to the boundary conditions of some correlation functions [21].@@@

We can look upon the above phenomenon as a spontaneous break–down of the up–down symmetry: the Hamiltonian of the model is symmetric, in a

zero field, with respect to spin reversal if one neglects the boundary terms; the phase transition manifests itself in the fact that there are equilibrium states in which the symmetry is violated only on the boundary and which are not symmetric even in the limit when the boundary recedes to infinity.

§8. – Microscopic description of the pure phases.

The description of the phase transition presented in Sect. **7** can be made more precise from the physical point of view as well as from the mathematical point of view. A deep and physically clear description of the phenomenon is provided by the theorem below, which also makes precise some ideas familiar from the droplet model [22].

Assume that the boundary condition is the (+)-boundary condition and describe a spin configuration $\underline{\sigma} \in \mathcal{U}(\Lambda)$ by means of the associated closed disjoint polygons $(\gamma_1, \dots, \gamma_n)$.

We regard the ensemble $\mathcal{U}(\Lambda)$ as equipped with the probability distribution attributing to $\underline{\sigma} = (\gamma_1, \dots, \gamma_n)$ a probability proportional to (7.3).

Then the following theorem holds:

Theorem. If β is large enough there exist $C > 0$ and $\rho(\gamma) > 0$ with $\rho(\gamma) \leq e^{-2\beta J|\gamma|}$ and such that a spin configuration $\underline{\sigma}$ randomly chosen out of the ensemble $\mathcal{U}(\Lambda)$ will contain, with probability approaching 1 as $\Lambda \to \infty$, a number $K_{(\gamma)}(\underline{\sigma})$ of contours congruent to γ such that

$$|K_{(\gamma)}(\underline{\sigma}) - \rho(\gamma)\,|\Lambda|| \leq C\sqrt{|\Lambda|}\, e^{-\beta J|\gamma|} \tag{8.1}$$

and this relation holds simoultaneosly for all Γ's. (In three dimensions one would have $|\Lambda|^{2/3}$ insted of $\sqrt{|\Lambda|}$.)

It is clear that the above Theorem means that there are very few contours (and that the larger they are the smaller is, in absolute and relative value, their number). The inequality (8.1) also implies that for some $C(\beta)$ there are *no* contours with perimeter $|\gamma| > C(\beta) \log |\Lambda|$ (with probability approaching 1 as $\Lambda \to \infty$). Hence a typical spin configuration in the grand canonical ensemble with (+)–boundary conditions is such that the large majority of the spins is "positive" and, in the "sea" of positive spins, there are a few negative spins distributed in small and rare regions (in a number, however, still of order of $|\Lambda|$).

Another nice result which follows from the results of Sect. **7** and from some improvement [24] of them concerns the behaviour of the equation of state near the phase transition region at low (enough) temperatures.

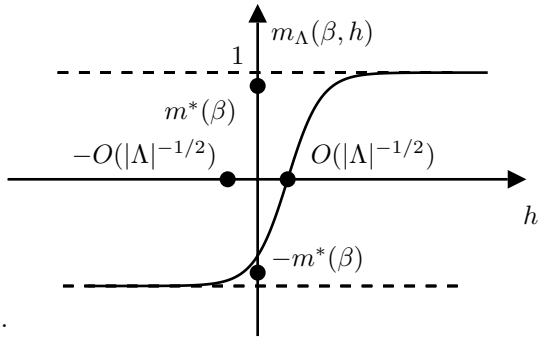


Fig. 4.

If Λ is finite the graph of $h \to m_\Lambda(\beta,h)$f will have a rather different behaviour depending on the possible boundary conditions; *e.g.* if the boundary condition is $(-)$ or $(+)$ one gets respectively the results depicted in Fig. 4 and 5.

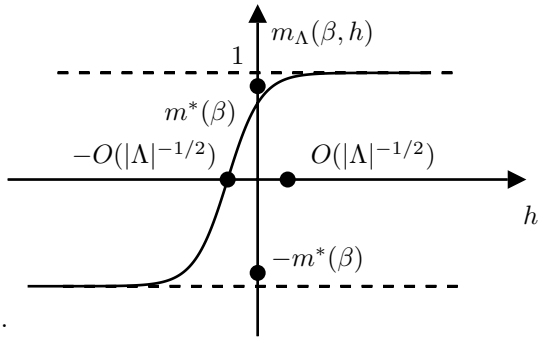


Fig. 5.

With periodic boundary conditions the diagram changes as in Fig. 6.

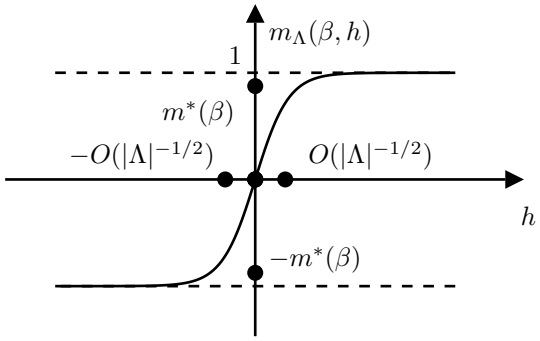


Fig. 6.

The thermodynamic limit $m(\beta,h) = \lim_{\Lambda\to\infty} m_\Lambda(\beta,h)$ exists for all $h \neq 0$

and the resulting graph is as shown in Fig. 7.

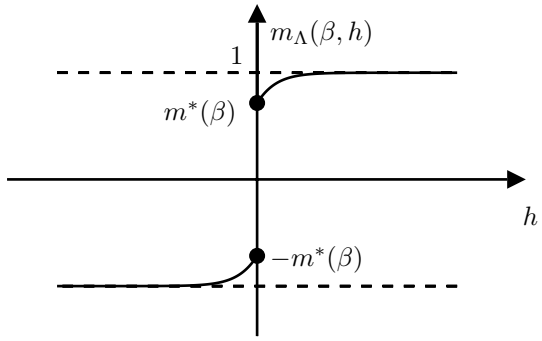


Fig. 7

At $h = 0$ the limit is not well defined and depends on the boundary condition (as it must). It can be proven, if β is large enough, that $\lim_{h\to 0^+} \frac{\partial\, m(\beta,h)}{\partial h} = \chi(\beta)$ is a finite number (*i.e.* the angle between the vertical part of the graph and the rest is sharp [24]).

The above considerations and results also furnish a clear idea of what a phase transition for a finite system means.

It is often stated that a finite systen "does not" show "sharp" phase transitions: however this statement is always made when considering one fixed boundary condition, usually of periodic or perfect-wall type. By taking into account the importance of the boundary terms we see which kind of phenomena occur in a finite system, if the corresponding infinite system has a sharp phase transition.

The next Section is devoted to the discussion of a number of problems concerning the generality of the definition of a phase transition as an instability with respect to the boundary perturbations, and other related problems. Note that an unpleasant limitation on the results discussed in this Section is the condition of low temperature ("β large enough").

§9. – Results on phase transitions in a wider range of temperature.

The results of the preceding Sections show that, at a low enough temperature, the Ising model is unstable with respect to changes in the boundary conditions. A natural question is whether one can go beyond the low temperature region and fully describe the phenomena in the region where the instability takes place. In the particular case of 2 dimensions it would also be natural to ask whether the maximum value of β to which an instability is associated is the one given by (6.9) which corresponds to the value of β

where the infinite volume free energy $f(\beta)$ has a singularity.

The above types of questions are very difficult and are essentially related to the already mentioned theory of the phase transitions based on the search and study of analytic singularities of the thermodynamic functions (which is a theory, however, that has still to be really developed).

Nevertheless a number of interesting partial results are known, which considerably improve the picture of the phenomenon of the phase transitions emerging from the previous Sections. A list of such results follows:

1) It can be shown that the zeros of the polynomial in $z = e^{\beta h}$ given by the given by the product of $z^{|\Lambda|}$ times the partition function (2.4) with periodic or perfect–wall boundary conditions lie on the unit circle: $|z| = 1$. It is easy to deduce, with the aid of Vitali's convergence theorem for equibounded analytic functions, that this implies that the only singularities of $f(\beta, h)$ in the region $0 < \beta < \infty$, $-\infty < h < +\infty$ can be found at $h = 0$.

A singularity appears if and only if the point $z = 1$ is an accumulation point of the limiting distribution (as $\Lambda \to \infty$) of the zeros on the unit circle. In fact if the zeros in question are $z_1, \ldots, z_{2|\Lambda|}$

$$\frac{1}{|\Lambda|} \log z^{|\Lambda|} Z(\beta, h, \Lambda, \text{periodic}) = 2\beta J + \beta h + \frac{1}{|\Lambda|} \sum_{i=1}^{2|\Lambda|} \log(z - z_i) \qquad (9.1)$$

and if $|\Lambda|^{-1}\cdot$(number of zeros of the form $z_j = e^{i\vartheta_j}$ with $\vartheta \le \vartheta_j \le \vartheta + d\vartheta) \xrightarrow[\Lambda\to\infty]{} \rho_\beta(\vartheta)\frac{d\vartheta}{2\pi}$ in a suitbale sense, we get from (9.1),

$$\beta f(\beta, h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log(z - e^{i\vartheta})\, \rho_\beta(\vartheta) d\vartheta + 2\beta J + \beta h \qquad (9.2)$$

where the last term comes from the $|z|^{|\Lambda|}$ appearing in (9.1).

The existence of the measure $\rho_\beta(\vartheta)\frac{d\vartheta}{2\pi}$ such that (9.2) is true follows, after some thought, from the existence of the thermodynamic limit $\lim_{\Lambda\to\infty} f_\Lambda(\beta, h) = f(\beta, h)$ [25].

2) It can be shown that the zeros of the partition function do not move too much under small perturbations of the spin–spin potential even if one allows "many spin" interactions; *i.e.* even if one perturbs the Hamiltonian (2.1) with perfect–wall boundary conditions into

$$\begin{aligned} H'_\Lambda(\underline{\sigma}) &= H_\Lambda(\underline{\sigma}) + (\delta H_\Lambda)(\underline{\sigma}) \\ (\delta H_\Lambda)(\underline{\sigma}) &= \sum_{k\ge 1} \sum_{x_1,\ldots,x_k \in \Lambda} \frac{1}{k!} J'(x_1, \ldots, x_k)\, \sigma_{x_1} \ldots \sigma_{x_k} \end{aligned} \qquad (9.3)$$

where $J'(X)$ is a function of the set $X = (x_1, \ldots, x_k)$ such that

$$||J'|| = \sup_{y\in Z^0} \sum_{y\in X} |J'(X)| \qquad (9.4)$$

is small enough.

More precisely, suppose that one knows that, when $J' = 0$, the zeros of the partition function in the variable $z = e^{\beta h}$ lie in a certain closed set N of the z–plane. Then if $J' \neq 0$ they lie in a closed set N^1 contained in a neighborhood of N which can be made as small as we please when $||J'|| \to 0$.

This result allows us to make a connection between the analyticity properties and the boundary condition instability as described in 3) below, [26].

3) There can be a boundary condition instability only in zero field and, in this case, if and only if the spectrum $\rho_\beta(\vartheta)$ has no gap around $\vartheta = 0$.

The proof of this result relies upon 2) and the remark that the correlation functions are functional derivatives with respect to $J'(x_1, \ldots, x_k)$ of the free energy defined by the Hamiltonian (9.3) [26].

4) Another question is whether the boundary condition instability is always revealed by the one–spin correlation function (as in Sect. **7**) or whether it might be shown only by some correlation functions of higher order. This question is answered by the following result.

There can be a boundary condition instability (at $h = 0$ and β fixed) if and only if

$$\lim_{h \to 0^-} m(\beta, h) \neq \lim_{h \to 0^+} m(\beta, h) \tag{9.5}$$

Note that, in view of what was said above (point 3)), $m(\beta, h) = \lim_{\Lambda \to \infty} m_\Lambda(\beta, h)$ is boundary condition independent as long as $h \neq 0$.

In other words there is a boundary condition instability if and only if there is spontaneous magnetization. This rules out the possibility that the phase transition could manifest itself through an instability of some higher–order correlation function which, practically, might be unobservable from an experimental point of view [27].

5) Point 4) implies that a natural definition of the critical temperature T_c is to say that it is the least upper bound of the T's such that (9.5) is true ($T = \beta^{-1}$). It is clear that, at this temperature, the gap around $\vartheta = 0$ closes and the function $f(\beta, h)$ has a singularity at $h = 0$ for $\beta > \beta_c = T_c^{-1}$ it can in fact be proven that if (9.5) is true for a given β_0 then it is true for all $\beta > \beta_0$ [28].

6) The location of the singularities of $f(\beta, 0)$ as a function of β remains an open question, see however [28]. In particular the question of whether there is a singularity of $f(\beta, 0)$ at $\beta = \beta_c$ is open. This implies that, at least in principle, it is still unproven that the singularity of the Onsager solution of the two–dimensional Ising model takes place at the critical point as defined in 5). It is, however, clear from the above considerations and from the fact, proved in the reference quoted in [6], that for $\beta > \beta_{c,0}$ (9.5)

certainly holds, that $\beta_{c,0} \geq \beta_c$ (see (6.9)).[1]

7) Finally another interesting question can be raised. For $\beta < \beta_c$ we have instability with respect to the boundary conditions (see 6) above): how strong is this instability? In other words, how many "pure" phases can exist?

Our intuition, in the case of the Ising model, suggests that there should be only two phases: the positively magnetized and the negatively magnetized ones.

To answer to the above question in a precise way it is necessary to agree on what a pure phase is [29]. We shall call an equilibrium state a "pure phase" if it is translationally invariant and if the correlation functions have a cluster property of the form

$$\langle \sigma_{x_1} \dots \sigma_{x_n} \sigma_{y_1+a} \dots \sigma_{y_m+a} \rangle \xrightarrow[a\to\infty]{} \langle \sigma_{x_1} \dots \sigma_{x_n} \rangle \langle \sigma_{y_1} \dots \sigma_{y_m} \rangle \tag{9.6}$$

where the convergence is understood in a very weak sense, *i.e.* the weakest sense which still allows us to deduce that the fluctuations of the extensive quantities are $o(|\Lambda|)$ [28], *i.e.*

$$\frac{1}{|\Lambda|} \sum_{a\in\Lambda} \langle \sigma_{x_1} \dots \sigma_{x_n} \sigma_{y_1+a} \dots \sigma_{y_m+a} \rangle \xrightarrow[\Lambda\to\infty]{} \langle \sigma_{x_1} \dots \sigma_{x_n} \rangle \langle \sigma_{y_1} \dots \sigma_{y_m} \rangle \tag{9.7}$$

i.e. the convergence in (9.6) takes place in the Cesaro–limit sense.

It can be proved that, in the case of the Ising model, the two states obtained as limits for $\Lambda \to \infty$ of finite volume states (cfr. Sect. **3**) corresponding to $(+)$–boundary conditions or $(-)$–boundary conditions are different for $\beta > \beta_c$ and are pure phases in the sense of (9.7) above [30].

Actually it can be proved that, in this case, the limits (9.6) exist in the ordinary sense [30] rather than in the Cesaro sense, and that at low temperature they are approached exponentially fast [31].

Furthermore, if β is large enough (*e.g.* in two dimension 10% larger than β_c), these two pure phases exhaust the set of pure phases [32]. For β close to β_c, however, the question is still open.[2]

[1] The identity $\beta_c \equiv \beta_{c,0}$ for the two–dimensional Ising model was later proved in [BGJS] and, independently, in [AM].

[2] It has been completely solved, since, in a remarkable series of papers based on the key work [Ru]: see [Hi],[Ai]. The work [Ru] has provided a real breakthrough and a lot of new ideas for the theory of the Ising model and percolation theory. The solution of this problem has led to the introduction of many new ideas and techniques in Statistical Mechanics and Probability Theory.

Having discussed the exact results about the structure of the phase transition and the nature of the pure phases, we shall turn in the nect Section to the phenomenon of coexistence of two pure phases.

§10. – Separation and coexistence of pure phases. Phenomenological considerations.

Our intuition about the phenomena connected with the classical phase transitions is usually based on the properties of the liquid–gas phase transition; this transition is experimentally investigated in situations in which the total number of particles is fixed (canonical ensemble) and in the presence of an external field (gravity).

The importance of these experimental conditions is obvious; the external field produces a non translationally invariant situation and the corresponding separation of the two phases. The fact that the number of particles is fixed determines, on the other hand, the fraction of volume occupied by the two phases. The phenomenon of phase transitions in the absence of an external field will be briefly discussed in Sect. **10**.

In the frame of the Ising model it will be convenient to discuss the phenomenon of the phase coexistence in the analogue of the canonical ensemble $\mathcal{U}(\Lambda, m)$, introduced and discussed in Sect. **2**, where the total magnetization $M = m|\Lambda|$ is held fixed.

To put ourselves in the phase transition region we shall take β large enough and:

$$m = \alpha\, m^*(\beta) + (1-\alpha)\,(-m^*(\beta)) = (1-2\alpha)\, m^*(\beta) \tag{10.1}$$

i.e. we put ourselves in the vertical "plateau" of the diagram $(m, h)_\beta$ (see Fig. 7).

Having fixed m as in (10.1) does not yet determine the separation of the phases in two different regions; to obtain this effect it will be necessary to introduce some external cause favouring the occupation of a part of the volume by a single phase. Such an asymmetry can be obtained at least in two ways: through a weak uniform external field (in complete analogy with the gravitational field of the liquid–vapour transition) or through an asymmetric field acting only on the boundary spins. This second way should have the same qualitative effect as the former, because in a phase transition region a boundary perturbation produces volume effects (this last phenomenon, which has been investigated in the previous Sections, is often also referred to as the "long range order" of the correlations).

From the mathematical point of view it is simpler to use a boundary asymmetry to produce a phase separations since it corresponds to a break–down of the up–down symmetry due only to the boundary spins (whose

number is relatively small).

To obtain a further, but not really essential, simplification of the problem consider the two–dimewnsional Ising model with $(+,-)$–cylindrical or $(+,+)$–cylindrical boundary conditions.

The spins adjacent to the bases of Λ act as symmetry–breaking external fields. The $(+,+)$–cylindrical boundary condition should, clearly, favour the formation inside Λ of the positively magnetized phase; therefore it will be natural to consider, in the canonical ensemble, this boundary condition only in the case that the total magnetization is fixed to be $+m^*(\beta)$ (see Fig. 7).

On the other hand the boundary condition $(+,-)$ favours the separation of phases (positively magnetized phase near the top of Λ and negatively magnetized phase near the bottom).

Therefore it will be natural to consider this boundary condition in the case of a canonical ensemble with magnetization $m = (1-2\alpha)\, m^*(\beta)$ (cfr. (10.1)).

In this last case one expects, as already mentioned, the positive phase to adhere to the top of Λ and to extend, in some sense to be discussed, up to a distance $\propto L$ from it; and then to change into the negatively magnetized pure phase.

To make precise the above phenomenological description we shall describe the spin configurations $\underline{\sigma} \in \mathcal{U}(\Lambda, m)$ through the associated sets of disjoint polygons (cfr. Sect. **6**).

Fix the boundary conditions to be $(+,+)$ or $(+,-)$–cylindrical boundary conditions and note that the polygons associated with a spin configuration $\underline{\sigma} \in \mathcal{U}(\Lambda, m)$ are *all closed* and of two types: the ones of the first type, denoted $\gamma_1, \dots, \gamma_n$, are polygons which do not encircle Λ, the second type of polygons, denoted by the symbols λ_α, are the ones which wind up around Λ.

So a spin configuration $\underline{\sigma}$ will be described by a set of polygons $(\gamma_1, \dots, \gamma_n, \lambda_1, \dots, \lambda_h)$. It is, perhaps, useful to remark once more that the configuration $\underline{\sigma}$ will be described by different sets of polygons according to which boundary condition is used. However, for a fixed boundary condition, the correspondence between spin configuration and sets of disjoint closed contours is one–to–one and the statistical weights of a configuration $\underline{\sigma} = (\gamma_1, \dots, \gamma_n, \lambda_1, \dots, \lambda_h)$ is (cfr. (6.4))

$$e^{-2\beta J(\sum_i |\gamma_i| + \sum_j |\lambda_j|)}$$

It should also be remarked that the above notation is not coherent with the notation of Sect. **6**, where the symbol λ is used for open polygons (absent here); we hope that this will not cause any confusion. The reason why we call λ the contours that go around the cylinder Λ is that they "look

like" open contours if one forgets that the opposite sides of Λ have to be identified.

It is very important to remark that if we consider the $(+,-)$ boundary conditions then the number of polygons of λ–type must be *odd* (hence $\neq 0$), while if we consider the $(+,+)$–boundary condition then the number of λ–type polygons must be *even* (hence it could be 0).

§11. – Separation and coexistence of phases. Results.

Bearing in mind the geometric description of the spin configuration in the canonical ensembles considered with the $(+,+)$–cylindrical or the $(+,-)$–cylindrical boundary conditions (which we shall denote briefly as $\mathcal{U}^{++}(\Lambda, m)$, $\mathcal{U}^{+-}(\Lambda, m)$) we can formulate the following Theorem [33]:

Theorem. For $0 < \alpha < 1$ fixed let $m = (1-2\alpha)\, m^*(\beta)$; then for β large enough a spin configuration $\underline{\sigma} = (\gamma_1, \dots, \gamma_n, \lambda_1, \dots, \lambda_{2h+1})$ randomly chosen out of $\mathcal{U}^{+-}(\Lambda, m)$ enjoys the properties 1)÷4) below with a probability (in $\mathcal{U}^{+-}(\Lambda, m)$) approaching 1 as $\Lambda \to \infty$:

1) $\underline{\sigma}$ contains *only one* contour of λ–type and

$$| \, |\lambda| - (1 + \varepsilon(\beta))L| < o(L) \tag{11.1}$$

where $\varepsilon(\beta) > 0$ is a suitable (α–independent) function of β tending to zero exponentially fast as $\beta \to \infty$.

2) If $\Lambda_\lambda, \Lambda'_\lambda$ denote the regions above and below λ we have

$$| \, |\Lambda_\lambda| - \alpha \, |\Lambda| \, | < \kappa(\beta) \, |\Lambda|^{3/4} \tag{11.2}$$

$$| \, |\Lambda'_\lambda| - (1-\alpha)|\Lambda| \, | <, \kappa(\beta) \, |\Lambda|^{3/4} \tag{11.3}$$

where $\kappa(\beta) \xrightarrow[\beta \to \infty]{} 0$ exponentially fast.

3) If $M_\lambda = \sum_{x \in \Lambda_\lambda} \sigma_x$, we have

$$|M_\lambda - \alpha m^*(\beta)|\Lambda|| < \kappa(\beta)|\Lambda|^{3/4} \tag{11.4}$$

and a similar inequality holds for $M'_\lambda = \sum_{x \in \Lambda'_\lambda} \sigma_x = m|\Lambda| - M_\lambda$.

4) If $K^\lambda_\gamma(\underline{\sigma})$ denotes the number of contours congruent to a given γ and lying in Λ_λ then, simoultaneously for all the shapes of γ:

$$| \, K^\lambda_\gamma(\underline{\sigma}) - \rho(\gamma)\alpha|\Lambda| \, | \leq C e^{-\beta J|\gamma|} \sqrt{|\Lambda|} \qquad C > 0 \tag{11.5}$$

where $\rho(\gamma) \leq e^{-2\beta J|\gamma|}$ is the same one as the one in the text of the theorem of Sect. **8**. A similar result holds for the contours below λ (cfr. the

comments on (8.1)).

It is clear that the above theorem not only provides a detailed and rather satisfactory description of the phenomenon of phase separation, but also furnishes a precise microscopic definition of the line of separation between the two phases, which should be naturally identified with the (random) line λ.

A very similar result holds in the ensemble $\mathcal{U}^{++}(\Lambda, m^*(\beta))$: in this case 1) is replaced by

1') *no* λ–type polygon is present

while 2),3) become superfluous and 4) is modified in the obvious way. In other words a typical configuration in the ensemble $\mathcal{U}^{++}(\Lambda, m^*(\beta))$ has the same appearance as a typical configuration of the grand canonical ensemble $\mathcal{U}(\Lambda)$ with (+)–boundary condition (which is described by the Theorem of Sect. **8**).

We conclude this Section with a remark about the condition that $0 < \alpha < 1$ has to be fixed beforehand in formulating the above Theorem. Actually the results of the theorem hold at fixed β (small enough) for all the α's such that $\varepsilon(\beta) < \min(\alpha, 1-\alpha)$, *i.e.* such that the line λ cannot touch the bases of Λ (in which case there would be additional physical phenomena).

§12. – Surface tension in two dimensions. Alternative description of the separation phenomena.

A remarkable application of the above theorem is the possibility of giving a microscopic definition of surface tension between the two pure phases [34].

We have seen that the partition functions

$$Z^{++}(\Lambda, \beta) = \sum_{\underline{\sigma} \in \mathcal{U}^{++}(\Lambda, m^*(\beta))} e^{-2\beta J(\sum_i |\gamma_i| + \sum_j |\lambda_j|)} \tag{12.1}$$

and (if $m = (1-2\alpha)m^*(\beta)$, $0 < \alpha < 1$)

$$Z^{+-}(\Lambda, \beta) = \sum_{\underline{\sigma} \in \mathcal{U}^{+-}(\Lambda, m)} e^{-2\beta J(\sum_i |\gamma_i| + \sum_j |\lambda_j|)} \tag{12.2}$$

will essentially differ, at low temperature, only because of the line λ (present in $\mathcal{U}^{+-}(\Lambda, m)$ and absent in $\mathcal{U}^{++}(\Lambda, m^*(\beta))$, see the preceding Section).

A natural definition (in two dimensions) of surface tension between the phases, based on obvious physical considerations, can therefore be given in terms of the different asymptotic behaviour of $Z^{++}(\Lambda, m^*(\beta))$ (or of the grand canonical $Z^{++}(\Lambda, \beta)$) and $Z^{+-}(\Lambda, m)$:

$$\tau(\beta) = \lim_{\Lambda\to\infty} \frac{1}{L} \log \frac{Z^{+-}(\Lambda, m)}{Z^{++}(\Lambda, m^*(\beta))} \tag{12.3}$$

The above limit (which should be α–independent for $\varepsilon(\beta) < \min(\alpha, 1-\alpha)$, cfr. the concluding remarks of the preceding Section) can be exactly computed at low enough temperature and it is given by

$$\tau(\beta) = -2\beta J - \log \tanh \beta J \tag{12.4}$$

which is the value computed by Onsager [3] by using a different definition, not based on the above detailed microscopic description of the phases and of the line of separation [35].

We conclude this Section with a brief discussion on one particular but very convenient alternative way of investigating the phenomenon of coexistence of two phases. Another still different way of investigating the phenomenon will be discussed in Sect. **14**.

Consider the grand canonical ensemble, but impose the following boundary conditions: the spins adjacent to the upper half of the boundary of Λ are fixed to be $+1$, while the ones adjacent to the lower half are -1 (and no periodicity condition). This is a $\underline{\varepsilon}$–type boundary condition (see Sect. **3**) which we shall denote $\mathcal{U}_0^{+-}(\Lambda)$.

It is clear that a configuration $\underline{\sigma} \in \mathcal{U}_0^{+-}(\Lambda)$ is described, under the above boundary condition, by one open polygon λ (surface in 3 dimensions) going from one side of Λ to the opposite side, and by a set of disjoint closed polygons (polyhedra in 3 dimensions) $(\gamma_1, \dots, \gamma_n)$.

The surface λ plays now the role of the polygons encircling Λ in the case of cylindrical boundary conditions (and 2 dimensions) and it is also clear that a theorem very similar to the ones already discussed should hold in this case. The above point of view is more relevant in the three–dimensional case where a "cylindrical" boundary condition would have a less clear physical meaning, and it would rather look as a mathematical device.

In the three dimensional case λ is a "surface" with a boundary formed by the square on $\partial\Lambda$ where is located the "break" between the spins fixed to be $+1$ and the ones fixed to be -1.

In the next section we investigate in more detail the structure of the line or surface of of separation between the phases.

§13. – The structure of the line of separation. What a straight line really is.

The Theorem of Sect. **11** tells us that, if β is large enough, then the line λ is almost straight (since $\varepsilon(\beta)$ is small). It is a natural question

to ask whether the line λ is straight in the following sense: suppose that λ, regarded as a polygon belonging to a configuration $\underline{\sigma} \in \mathcal{U}^{(+-)}(\Lambda, m)$ (cfr. Sect. **11**), passes through a point $q \in \Lambda$; then we shall say that λ is "straight" or "rigid" if the (conditional) probability $\mathcal{P}_\lambda$ that λ passes also through the site q', opposite[3] to q on the cylinder Λ, *does not* tend to zero as $\Lambda \to \infty$, otherwise we shall say that λ is *not rigid* or *fluctuates*. Of course the above probabilities must be computed in the ensemble $\mathcal{U}^{+-}(\Lambda, m)$.

Alternatively (and essentially equivalently) we can consider the ensemble $\mathcal{U}_0^{+-}(\Lambda)$ (see Sect. **12**, *i.e.* the grand canonical ensemble with the boundary condition with the boundary spins set to $+1$ in the upper half of Λ, vertical sites included, and to -1 in the lower half). We say that λ is rigid if the probability that λ passes through the center of the box Λ (*i.e.* 0) *does not* tend to 0 as $\Lambda \to \infty$; not rigid otherwise.

It is rather clear that what the above notion of rigidity means: the "excess" length $\varepsilon(\beta)L$, see (11.1), can be obtained in two ways: either the line λ is essentially straight (in the geometric sense) with a few "bumps" distributed with a density of order $\varepsilon(\beta)$ or, otherwise, the line λ is bent and, therefore, only locally straight and part of the excess length is gained through the bending.

In three dimensions a similar phenomenon is possible. As remarked at the end of the last Section, in the ensemble $\mathcal{U}_0^{+-}(\Lambda)$, in this case λ becomes a surface with a square boundary fixed at a certain height (*i.e.* 0), and we ask whether the centre of the square belongs to λ with non vanishing probability in the limit $\Lambda \to \infty$.

The rigidity or not of λ can, in principle, be investigated by optical means; one can have interference of coherent light scattered by surface elements of λ separated by a macroscopic distance only if λ is rigid in the above sense.

It has been rigorously proven that, at least at low temperature, the line of separation λ is *not* rigid in 2 dimensions (and the fluctuation of the middle point is of the order $O(\sqrt{|\Lambda|})$). On the contrary, in 3 dimensions it has been shown that the surface λ is rigid at low enough temperature [36] and it is conjectured to become non rigid at higer temperatures (still below the critical point, of course).[4]

An interesting question remains open in the three dimensional case and is the following: it is conceivable that the surface, although rigid at low temperature, might become loose at a temperature $\tilde{T}_c$ smaller than the critical temperature T_c (defined as the pargest temperature below which there are at least two pure phases).

It would be interesting to examine the available experimental data on the structure of the surface of separation to set limits on $T_c - \tilde{T}_c$ in the case of

[3] *i.e.* on the same horizontal line and $L/2$ sites apart.

[4] This is now known, see footnote in Sect. **9**.

the liquid–gas phase transition where such a phenomenon can conceivably occurr even though a theory of it is far from being in sight, at least if one requires a degree of rigour comparable to that displayed in the treatment of the results so far given for the Ising model.

We conclude by remarking that the rigidity of λ is connected with the existence of translationally noninvariant equilibrium states (see Sect. **3**).

It seems almost certain that, in 2 dimensions, because of the discussed non rigidity of λ there are no translationally noninvariant states [36].

Note that the existence of translationally noninvariant equilibrium states is not necessary for the description of the coexistence phenomena. The theory of the two dimensional Ising model developed in the preceding sections is a clear proof of such a statement [36].

14. – Phase separation phenomena and boundary conditions. Further results.

The phenomenon of phase separation described in Sect. **10** and **11** is the ferromagnetic analogue of the phase separation between a liquid and a vapour in the presence of the gravitational field.

It is relevant to ask to what extent an external field (or some equivalent boundary condition) is really necessary; for instance one could imagine a situation in which two phases coexist in the absence of any external field.

Let us discuss first some phenomenological aspects of the liquid–gas phase separation in the absence of external fields. One imagines that, if the density is fixed and corresponds to some value on the "plateau" of the phase diagram, then the space will be filled by vapour and drops of liquid in equlibrium. Note that the drops will move and, from time to time, collide; since the surface tension is negative the drops will tend to cluster togheter and, eventually, in an equilibrium situation there will be just one big drop (*i.e.* the drop surface will be minimal). The location of the drop in the box Λ will depend on how the walls are made and how they interact with the particles within Λ.

Let us consider some extreme cases:

1) the walls "repel" the drops,

2) the walls "attract" the drops,

3) the wall is perfect and does not distinguish bewtween the vapour and the liquid.

In the first case the drop will stay away from the boundary $\partial\Lambda$ of Λ. In the second case the drop will spread on the walls, which will be wet as much as possible. In the third case it will not matter where the drop is; the drop will be located in a position that minimizes the "free" part of its boundary (*i.e.* the part of the boundary of the drop not on $\partial\Lambda$). This means that the

drop will prefer to stay near a corner rather than wetting all the wall.

Let us translate the above opicture into the Ising model case. Assume that β is large and $m = (1-2\alpha)m^*(\beta)$ (see Fig. 7) (*i.e.* assume that the magnetization is on the vertical plateau of the $(m,h)_\beta$ diagram in Fig. 7).

Then the conditions 1), 2), 3) can be realized as follows:

1) The spins adjacent to the boundary are all fixed to be $+1$. This favours the adherence to the boundary of the positively magnetized phase.

2) The spins adjacent to the boundary are all fixed to be -1. This favours the adherence to the boundary of the negatively magnetized phase.

3) There are no spins adjacent to the boundary, *i.e.* we consider *perfect wall* boundary conditions (see Sect. **3**).

The rigorous results available in the case of the Ising model confirm the phenomenological analysis based on the liquid–vapour coexistence [23]:

Theorem. Fix $0 < \alpha < 1$ and consider $(-)$–boundary conditions. The a spin configuration $\underline{\sigma}$ randomly extracted from the canonical ensemble with magnetization $m = (1-2\alpha)m^*(\beta)$ has, if β is large enough, properties 1)÷3) below with a probability tending to 1 as $\Lambda \to \infty$.

1) There is *only one* γ such that $|\gamma| > \frac{1}{333}\log|\Lambda|$ and it has the property[5]

$$\big|\,|\gamma| - 4\sqrt{(1-\alpha)|\Lambda|}\,\big| \le \delta(\beta)\sqrt{|\Lambda|} \tag{14.1}$$

with $\delta(\beta) \to 0$ as $\beta \to \infty$ (exponentially fast);

2) The area enclosed by γ is $\vartheta(\gamma)$:

$$\big|\,\vartheta(\gamma) - (1-\alpha)|\Lambda|\,\big| \le \kappa(\beta)\,|\Lambda|^{3/4} \tag{14.2}$$

3) The magnetization $M(\vartheta(\gamma))$ inside γ is on the average equal to $-m^*(\beta)$ and, more precisely,

$$\big|\,M(\vartheta(\gamma)) + m^*(\beta)\,(1-\alpha)\,|\Lambda|\,\big| \le \kappa(\beta)\,|\Lambda|^{3/4} \tag{14.3}$$

and, therefore, the average magnetization outside $\vartheta(\gamma)$ is $+m^*(\beta)$.

This Theorem holds also in 3 dimensions but the exponent of $|\Lambda|$ in (14.1) changes (from $\frac{1}{2}$ to $\frac{2}{3}$).

[5] The number 333 is just an arbitrary constant and it is reported here because it appeared in the original literature [31], as a joke referring to the papers on KAM theorem ("Arnold's constant"). In fact it looks today somewhat confusing and quite strange: the modern generations do not seem to appreciate this kind of humour any more; they became more demanding and would rather ask here for the "best" constant; this is my case as well.

The above Theorem shows that a typical configuration consists of a positively magnetized pure phase adherent to the boundary and of a "drop" of negatively magnetized phase *not adhering* to the boundary (since γ is closed). The size of the drop is $\sim \sqrt{(1-\alpha)|\Lambda|}$ (as it should be).

Note that the drop is almost square in shape (as follows from (14.1),(14.2)): this should not be astonishing since the space is discrete and the isoperimetric problem on a square lattice has the square as a solution (rather than a circle).

The opposite situation is realized if one fixes a $(-)$–boundary condition; a square drop forms in the middle of the box with side $\sim \sqrt{\alpha|\Lambda|}$ and average magnetization $-m^*(\beta)$.

Finally if the boundary condition is of perfect wall type ($\mathcal{B}_\Lambda(\underline{\sigma}) \equiv 0$), then the above Theorem does not hold and one can expect to prove (say, in 2 dimensions) that a typical spin configuration has just one open contour λ (with ends on $\partial\Lambda$) which separates the space in two parts which are occupied by the opposite phases; the line λ should be the shortest possible compatibly with the condition that the volume Λ is divided by it into two regions of volume essentially $\alpha|\Lambda|$ and $(1-\alpha)|\Lambda|$ (respectively occupied by the positively magnetized phase and by the negatively magnetized phase). The results just described for the case $\mathcal{B}_\Lambda(\underline{\sigma}) \equiv 0$ have never been proved though there is evidence for their truth [32].

If one interprets the spins equal to $+1$ as particles and the spins equal to -1 as empty sites, then one has a lattice gas model which undergoes a liquid–vapour phase transition which presents the phenomenological aspects outlined at the beginning of this Section for these transitions.

To conclude we remark that, in the phase separation phenomenon, the finiteness of the box only plays the role of fixing the density and keeping the vapour tension. The detailed structure of the phenomenon depends on the boundary conditions which, in experimental situations, turn out to be something intermediate between the thre extreme cases discussed above.

Note that (14.1) does not provide a satisfactory estimate of $|\lambda|$ since the allowed error is still of the order of $\sqrt{|\Lambda|}$; it is an open problem to obtain a better estimate of $|\lambda|$ of the type (11.1) (*i.e.* with an error much smaller than $O(\sqrt{|\Lambda|})$). It is also an open problem to find an expression for the surface tension of the square drop (which is expected to be the same as (12.3) in two dimensions); see the Introduction to [34].[6] A third problem is the investigation of the dependence of the correlation functions on the distance from the surface of the drop.

[6] These problems have been solved in a series of subsequent papers [CDR], [DKS], [Pf] (emozionante!), see also [M], [KM]. The last results that I expected turned out to be different from my naive expectation above (accidenti!).

The analogues of the first two questions just raised have been satisfactorily answered in the 2–dimensional Ising model with cylindrical boundary conditions (see Sect. **11** and **12**), *i.e.* in the case of an "infinite" drop with a flat surface.

The third problem has been only approximately studied even in the case of a flat drop [37].

§15. – Conclusions and open problems.

In the preceding Section we have dealt with the case of a nearest neighbour Ising model. It has become customary, in the literature, to call with the name of Ising model more general models in which the "bulk" (*i.e.* without the boundary interactions and conditions) Hamiltonian has the form

$$-h\sum_{x_i}\sigma_{x_i}-\sum_{i<j}J_2(x_i,x_j)\sigma_{x_i}\sigma_{x_j}-\sum_{i<j<k}J_3(x_i,x_j,x_k)\sigma_{x_i}\sigma_{x_j}\sigma_{x_k}+\dots \tag{15.1}$$

where the potentials $J_n(x_1,\dots,x_n)$ are translationally invariant functions of $(x_1,\dots,x_n)$ and satisfy certain restrictions of the type:

$$\sum_x|J_2(0,x)|+\sum_{x,y}|J_3(0,x,y)|+\dots<+\infty \tag{15.2}$$

If only pair potentials are present, *i.e.* if the bulk Hamiltonian has the form

$$h\sum_x\sigma_x+\sum_{i<j}J(x_i-x_j)\sigma_{x_i}\sigma_{x_j} \tag{15.3}$$

and if $J(r)\le 0$, then most of the results described in this paper and appropriately reformulated have either already been proved or are being proved or are very reasonable conjectures [38].

Many results shall stay true for more general pair potentials and for other models (like continous gases) at least from the qualitative point of view; in fact it is reasonable that the results selected here for discussion should have, at least qualitatively, an analogue in the "general" case of a classical (as opposed to quantum) phase transition.

Results such as analitycity and absence of phase transitions at high temperature, or exact solutions, are a peculiarity of the lattice models and have been, therefore, left out. I made some exceptions to the above rule of selection of results by quoting some of the exact results from Onsager's solution of the 2–dimensional Ising model.

Below I list a number of rather randomly chosen and interesting open problems suggested by the topics of this article.

1) It would be interesting to fill the gap between T_c and the maximum value ($\sim 10\%$ of T_c in 2 dimensions) for which one can prove that there are only two pure phases. This is related to other problems such as the conjectured identity, in 2 dimensions, of β_c and $\beta_{c,O}$ (see (6.9) and point 4) in Sect. **9**).[7]

2) The solution of the 2 dimensional Ising model is based on the so called "transfer matrix". The investigation of the transfer matrix has been pursued in some detail in the case of periodic or open boundary conditions in two or three dimensions [40], see also [3,41].
It would be of interest to study the transfer matrix with nonsymmetric boundary conditions. In particular it would be of interest to study the transfer matrix between two rows (or planes) where the line (or surface) of separation should pass (if straight). A qualitative difference should arise between two and three dimensions (see, for more details, the Appendix).

3) In Fig. 7 we see that the isotherm $m(\beta, h)$ as a function of $h > 0$ abruptly ends at $h = 0$. It is still an open question whether $h = 0$ is an analytic singularity of $m(\beta, h)$ or whether $m(\beta, h)$ can be analytically continued to $h < 0$. There is strong evidence of a singularity [42].[8]

4) In the case the answer to 3) is in agreement with the conjecture, how one can explain the metastability phenomena [43]?[9]

5) It would be of interest to find generalizations of the phase coexistence theory to other lattice models for which phase transitions are proven to take place [44].[10]

6) The existence of phase transitions has been recently proved, for the first time, for a continuous system. It would be of interest to analyze the phenomenon of the phase coexistence in this case [45].[11]

7) If for a sytem a phase transition is known to take place, when can one answer the question of how many pure phases exist?

8) A detailed description of the correlation functions near the line or surface of separation has still to be discussed (see [36,37]).

---

[7] These problems have been entirely solved later, see [Ru], [Hi], [Ai], [BGJS], [AM].

[8] Indeed it has been shown that at $h = 0$ there is an essential singularity, at least at large β, although the function $m(\beta, h)$ is infinitely differentiable as a function of h for $h \geq 0$ [I].

[9] The metastability has been investigated in great detail as a dynamical phenomenon and the results are very many and very varied, see for instance [CCO], [KO], [MOS], [Sc].

[10] This has been done for a great variety of models, [BLPO].

[11] This has been done [BLPO].

9) It would be of interest to investigate the microscopic definition of surface tension in the particular case of the 3–dimensional Ising model (which, so far, has not been studied).
10) It would be of interest to prove that, in 3 dimensions, the surface tension $\tau(\beta)$ is such that $\tau(\beta) + 2\beta J$ is analytic in $e^{-\beta J}$.
11) It would be of interest to investigate the phase transitions in models not showing the up-down symmetry like the ones obtained by choosing in (15.1) $J_3 \neq 0$ [46].[12]
12) Three more open problems are listed at the end of Sect. **14**.

APPENDIX

Transfer matrix in the Ising model.

Consider the one dimensional Ising model with periodic boundary conditions. If $\sigma_{L+1} \equiv \sigma_1$ the partition function $Z(\Lambda, \beta, h)$ can be written as:

$$\sum_{\sigma_1 \dots \sigma_L} \prod_{i=1}^{L} e^{\beta J \sigma_i \sigma_{i+1} + \beta h \sigma_i} = \sum_{\sigma_1 \dots \sigma_L} \prod_{i=1}^{L} e^{\frac{\beta}{2} h \sigma_i} e^{\beta J \sigma_i \sigma_{i+1}} e^{\frac{\beta}{2} h \sigma_{i+1}} = \\ = \sum_{\sigma_1 \dots \sigma_L} V_{\sigma_1 \sigma_2} V_{\sigma_2 \sigma_3} \dots V_{\sigma_L \sigma_1} = \operatorname{Tr} V^L \tag{a.1}$$

where V is a two-by-two matrix such that ($\sigma, \sigma' = \pm 1$):

$$V_{\sigma\sigma'} = e^{\frac{\beta}{2} h \sigma} e^{\beta J \sigma \sigma'} e^{\frac{\beta}{2} h \sigma'}, \qquad V = \begin{pmatrix} e^{\beta(h+J)} & e^{-\beta J} \\ e^{-\beta J} & e^{-\beta(h+J)} \end{pmatrix} \tag{a.2}$$

If $\lambda_+ > \lambda_-$ are the two eigenvalues of V, we find

$$Z(\Lambda, \beta, h) = \lambda_+^L + \lambda_-^L \tag{a.3}$$

$$\beta f(\beta, h) = \lim_{L \to \infty} \frac{1}{L} \log Z = \log \lambda_+ \tag{a.4}$$

It is easy to check that $\lambda_+(\beta, h)$ is analytic ib β and h for $0 < \beta < \infty$ and $-\infty < h < \infty$, *i.e.* there are no phase transitions (as singularities of $f(\beta, h)$.

A similar method can be applied to the two dimensional Ising model (Λ is now a $M \times N$ box). Suppose, for simplicity, $h = 0$, then $Z(\beta, h)$ is given by:

[12] This has been well understood at low temperature in the basic paper [PS] which generated a great number of rather complete studies of phase coexistence at low temperature.

$$\begin{aligned}&\sum_{\underline{\sigma}}\prod_{i=1}^{M}\prod_{j=1}^{M}e^{\beta J\sigma_{i,j}\sigma_{i+1,j}+\beta J\sigma_{i,j}\sigma_{i,j+1}}=\\&=\sum_{\underline{\sigma}_1}\cdots\sum_{\underline{\sigma}_M}\prod_{i=1}^{M}\Big\{\prod_{j=1}^{N}e^{\frac{\beta J}{2}\sigma_{i,j}\sigma_{i,j+1}}e^{\beta J\sigma_{i,j}\sigma_{i+1,j}+\frac{\beta J}{2}\sigma_{i+1,j}\sigma_{i+1,j+1}}\Big\}\end{aligned}\tag{a.5}$$

where in the second line we denote by $\underline{\sigma}_i=(\sigma_{i,1},\dots,\sigma_{i,N})$ all the spins on the i-th row of Λ; the periodic boundary conditions are imposed by setting $\underline{\sigma}_1\equiv\underline{\sigma}_{M+1}$ and $\sigma_{i,1}\equiv\sigma_{i,N+1}$.

Clearly, if we define the $2^N\times 2^N$ matrix

$$\begin{aligned}V_{\underline{\sigma},\underline{\sigma}'}&=\prod_{j=1}^{N}e^{\frac{\beta J}{2}\sigma_j\sigma_{j+1}}e^{\beta J\sigma_j\sigma'_j+\frac{\beta J}{2}\sigma'_j\sigma'_{j+1}}=\\&=\exp\sum_{i=1}^{N}\big(\frac{\beta J}{2}\sigma_j\sigma_{j+1}+\beta J\sigma_j\sigma'_j+\frac{\beta J}{2}\sigma'_j\sigma'_{j+1}\big)\end{aligned}\tag{a.6}$$

where $\sigma_1\equiv\sigma_{N+1}$, $\sigma'_1=\sigma'_{N+1}$, we realize that

$$Z(\Lambda,\beta)=\ \mathrm{Tr}\,V^M\tag{a.7}$$

We have dealt so far only with periodic boundary conditions. We could introduce transfer matrices also in the case of other boundary conditions.

For instance, assume, for simplicity, that there are periodic boundary conditions alog the columns; we shall consider the three cases below:

1) "perfect wall" boundary conditions along the rows;

2) boundary conditions on the rows corresponding to the existence, on the lattice sites adjacent to the end points of the rows, of fixed spins $\varepsilon_i=+1$ (or $\varepsilon_i=-1$) for all the i's;

3) boundary conditions which are of the same type as in 2) but *half* the rows end in positive spins (say the upper half) and half in a negative spin.

We shall now write down a transfer matrix expression for $Z(\Lambda,\beta)$ in the above cases. In case 1) $Z(\Lambda,\beta)=\mathrm{Tr}\,V^{(1)M}$ where:

$$V^{(1)}_{\underline{\sigma},\underline{\sigma}'}=e^{\sum_{j=1}^{N-1}(\frac{\beta J}{2}(\sigma_j\sigma_{j+1}+\sigma'_j\sigma'_{j+1})+\sum_{j=1}^{N}\beta J\sigma_j\sigma'_j)}\tag{a.8}$$

In case 2) $Z(\Lambda,\beta)=\mathrm{Tr}\,V^{(\pm)M}$ where:

$$V^{(\pm)}_{\underline{\sigma},\underline{\sigma}'}=e^{\pm\beta J(\sigma_1+\sigma'_1+\sigma_N+\sigma'_N)}V^{(1)}_{\underline{\sigma},\underline{\sigma}'}\tag{a.9}$$

In case 3), assuming here that the height of Λ is $M+1$ with M even, we have that $Z(\Lambda,\beta)=\ \mathrm{Tr}\,(V^{(+)})^{M/2}V^{(3)}(V^{(-)})^{M/2}$ with

$$V^{(3)}_{\underline{\sigma},\underline{\sigma}'} = e^{\pm\beta J(\sigma'_1+\sigma'_N-\sigma_1-\sigma_N)} V^{(1)}_{\underline{\sigma},\underline{\sigma}'} \qquad (a.10)$$

The transfer matrix V in (a.7) is the matrix that was diagonalized int he famous paper of Onsager [3]. The matrix $V^{(1)}$ has also been diagonalized exactly in ref. [47].

The matrices $V^{(\pm)}$ have, so far, never been studied; nor the $V^{(3)}$ has been studied.[13]

The problem of computing the partition function can be formulated similarly in the three dimensional case. Some very interesting results on the spectral properties of the generalization to dimension three of the matrix V (periodic boundary conditions) have been obtained in ref. [48].

In three dimensions one expects that the analogue of $V^{(3)}$ (in contrast to $V^{(1)}$, $V^{(\pm)}$) has spectral properties which radically differ from those of V. In two dimensions the phenomenon should not occurr and all the above matrices should have the same spectrum (asymptotically as $\Lambda \to \infty$). As mentioned in Sect. **15**, problem 2), this should be related to the fact that $V^{(3)}$ should contain some information about the rigidity of the line or surface of phase separation (which is "rigidly sitting" right near the two lines between which $V^{(3)}$ "transfers").

A very interesting heuristic analysis of the spin correlation functions in terms of the transfer matrix has been done in ref. [49]. The paper of ref. [48] (written independently of [49]) has been devoted to the attempt to understand completely this analysis.

NOTES and REFERENCES

[1] For a history of the Ising model see S.G. BRUSH: *Rev. Mod. Phys.*, **39**, 883 (1969).

[2] R. PEIERLS: *Proc. Cambridge Phil. Soc.*, **32**, 477 (1936).

[3] L. ONSAGER: *Phys. Rev.* **65**, 117 (1944).

[4] T.D. LEE and C.N. YANG: *Phys. Rev.*, **87**, 410 (1952).

[5] The mean field theory, as it appears from the literature quoted below, is a theory of phase transitions due to very-long-range and weak forces. The Ising model is, in contrast, a theory based on strong and short-ranged forces. The original approach to the van der Waals theory (also called mean field theory) can be found in the book by S. CHAPMAN and T. COWLING: *The mathematical Theory of Nonuniform Gases* (Cambbridge, 1953), p. 284A more refined and interesting formulation is in N.G. van Kampen:: *Phys. Rev.*, **135**A, 362 (1964). A precise and very clear theory is in J.L. LEBOWITZ and O. PENROSE: *J. Math. Phys.*, **7**, 98 (1966). The first precise understanding (and full

[13] The $V^{(\pm)}$ matrices have, since, been diagonalized in [AM]. Many more exact calculations of interesting quantities have been performed, [A].

proof in particular cases) of mean field theory comes from P. HEMMER, M. KAC and G.E. UHLENBECK in a series of papers appeared in *J. Math. Phys.* and reproduced, with introductory remarks, in E. LIEB and D.C. MATTIS: *Mathematical Physics in one dimension*, (New York, 1966). A more phenomenological but very interesting and original theory is in the book of R.H. BROUT: *Phase transitions* (New York, 1965), where the most common phase transitions are treated from the unifying point of view of the mean field theory.

[6] The original solution for the free energy of the Ising model in 2 dimensions can be found in [3]. It was preceded by the proof of existence in [2][14] and by the exact location of the critical temperature by H.A. KRAMERS and G.H. WANNIER: *Phys. Rev.*, **60**, 252 (1941). The spontaneous magnetization was found by L. ONSAGER : *Suppl. Nuovo Cimento*, **6**, 261 (1949), but the details were never published; it was subsequently rediscovered by C.N. YANG: *Phys. Rev.*, **85**, 809 (1952). A modern derivation of the solution is found in the review article of T.D. SCHULTZ, D.C. MATTIS and E. LIEB: *Rev. Mod. Phys.*, **36**, 856 (1964). Another interesting older review article is the paper by G.F. NEWELL and E.W. MONTROLL: *Rev. Mod. Phys.*, **25**, 353 (1953). A combinatorial solution has been found by M. KAC and M. WARD and can be found in the book by L. LANDAU and L.E. LIFSCHITZ: *Physique Statistique* (Moscow, 1967), p. 538. This derivation is not above criticism; an up-to-date paper on the subject is due to F.A. BEREZIN: *Russian Math. Surveys*, **24**, 1 (1969). Another approach to the solution (the "pfaffian method") can be found in P.W. KASTELEYN: *Physica*, **30**, 615 (1961).

[7] See the review paper by M.E. FISHER: *Rep. Progr. Theor. Phys.*, **30**, 615 (1967), pp. 677-702 and appended references.

[8] It is the case of $MnCl_2 \cdot 4H_2O$, see M.E. FISHER, M.F. SYKES: *Physica*, **28**, 939 (1962); see also [7].

[9] R.L. DOBRUSHIN: *Theory of probability and its applications*, **13**, 197 (1968); *Funct. Anal. and Appl.*, **2**, 292 (1968); **2**, 302 (1968); **3**, 22 (1968). F. SPITZER: *Am. Math. Monthly*, **78**, 142 (1971).

[10] D. RUELLE: *Statistical Mechanics*, (New York, 1969), p. 168, 161.

[11] R.B. GRIFFITHS: J. Math. Phys., **8**, 478 (1967); M.E. FISHER: *Phys. Rev.*, **162**, 475 (1967); H.S. GREEN and C.H. HURST: *Order Disorder Phenomena*, (New York, 1964).

[12] D.C. MATTIS: *Theory of Magnetism*, (London, 1965).

[13] This term is usually omitted and in some sense its importance has only recently been recognized after the works of R.L. DOBRUSHIN: see ref. [9], and of O. LANFORD and D. RUELLE: *Comm. Math. Phys.*, **13**, 194 (1969). It is one of the purposes of this article to emphasize the role of this term in the theory of phase transitions.

[14] K. HUANG: *Statistical Mechanics* (New York, 1963).

[15] Here we are cavalier about the problem of what $\partial/\partial m$ means, since $g_\Lambda(\beta.m)$ is only defined for certain rational values of m (whose number is finite). One could, for instance,

[14] The proof was carried further by van der Waerden, [Wa].

extend $g_\Lambda(\beta, m)$ to all m's by considering instead of $g_\Lambda(\beta, m)$ its convex envelope (or also one could prefer to consider the g_Λ obtained by linear interpolation from (2.6)). This is not very satisfactory but should not be really important for large ystems (it would be interesting to make more precise this point). With the above definition of g_Λ the function $m(h)$ might have jumps at finite Λ.

[16] This definition is inspired by O. LANFORD and D. RUELLE, cit. [13], where the equivalence of the above definition with a number of other possible definitions is shown. For instance the definition in question is equivalent to the one based on the requirement that the correlation functions should be a solution of the Kirkwood–Salsburg equations. It is also equivalent to the definition of equilibrium state in terms of tangent planes (*i.e.* functional derivatives of a suitable functional: see D. RUELLE, *Statistical Mechanics*, cit. [10], p. 184). It should be said that these proofs of equivalence are not always explicitly derived in the quoted paper by LANFORD and RUELLE; they are, however, an easy corollary of their results and appear, derived in detail, in the so far unpublished lecture notes of the lectures delivered by the author at the Courant Institute, September 1971, preprint.[15]

[17] The solution of the Ising model in zero field can be found in the paper by G. NEWELL and E.W. MONTROLL, cit. [6].

[18] The expansion can be used as a starting point for the combinatorial solution mentioned in [6]. See L. LANDAU and E.L. LIFSCHITZ, cit. [6].

[19] Of course we do not attach a deep physical meaning to the difference between these two approaches. Clearly they should be equival;ent if one pretended to extract all possible information from them. What is really important is that the first questions raised by both approaches are very interesting and relevant from a physical point of view. One of the goals of the analytic theory of phase transitions is to understand the nature of the singularity at the critical point and along the break of the isotherms. A lot of interest has been devoted to this point and a number of enlightening phenomenological results are avilable. However the number of complete results on the matter is rather limited. An idea of the type of problems that are of interest can be gotten by reading the papers of P.W. KASTELEYN: in *Fundamental Problems in Statistical Mechanics. II*, edited by E.G.D. COHEN (Amsterdam, 1968), or the more detailed paper by M.E. FISHER, cit. [7].

[20] This geometric picture of the spin configurations can be traced back at least as far as Peierls' paper , cit. [2], see also [Wa], and has been used, togheter with formula (4.11) to derive (6.8) ("Kramers–Wannier duality" relation)and (6.9) by H.A. KRAMERS and G.H. WANNIER, cit. [6]. A recent interesting genearlization of the duality concept has been given by F.J. WEGNER: *J. Math. Phys.*, **12**, 2259 (1971), where some very interesting applications can be found as well as references to earlier works. The duality relation between (+), or (−), boundary conditions and open boundary conditions (which is used here) has been developed in a conversation with G. BENETTIN, G. JONA-LASINIO, A. STELLA. The reader can find other interesting relations in their paper to

[15] I never published them: eventually part of them was included in my book [Ga2].

appear in *Lett. Nuovo Cimento*, (June 1972).[16]

[21] The above proof is due to R.B. GRIFFITHS and, independently, to R.L. DOBRUSHIN and it is inspired by R. PEIERLS, cit. [2], see also [Wa].

[22] M.E. FISHER: *Physics, Physica, Fizika*, **3**, 255, 1967.

[23] This theorem is due to R.A. MINLOS and J.G. SINAI: *Trans. Moscow Math. Soc.*, **19**, 121 (1968); *Math. USSR Sbornik*, **2**, 335 (1967).

[24] F.A. BEREZIN and J.G. SINAI: *Trans. Moscow Math. Soc.*, **19**, 219 (1967),

[25] Here the symbol $\rho_\beta(\vartheta)d\vartheta/2\pi$ has not to be taken too seriously; it really denotes a measure on the circle and this measure is not necessarily $d\vartheta$–continuous. Also the "convergence" statement really means the existence of a measure such that (6.2) holds for all real z. The original proof of this theorem is due to T.D. LEE and C.N. YANG, cit. [4]. A much stronger and general statement leading, inparticular, to Lee–Yang's theorem is in D. RUELLE: *Phys. Rev. Lett.*, **26**, 303 (1971). Ruelle's theorem has been the last of a series of improvements and generalizations of Lee-Yang's theorem; see references in Ruelle's paper.

[26] D. RUELLE: *The use of small external fields...*, preprint, to appear in *J. Math. Phys.*

[27] A. MARTIN–LÖF and J.L. LEBOWITZ: *Comm. Math. Phys.*, **25**, 276 (1972).

[28] R.B. GRIFFITHS: *J. Math. Phys.*, **8**, 478 (1967); M.E. FISHER: *Lectures in Physics*, vol. **7**C (Boulder, Colorado (1965)).

[29] The definition below is due to D. RUELLE: *Statistical Mechanics*, cit. [10], p. 161.

[30] This is an unpublished result of R.B. GRIFFITHS. His proof is reported in G. GALLAVOTTI, A. MARTIN-LÖF and S. MIRACLE–SOLÉ: in *Springer lecture notes in Physics*, vol. **20**, edited by A. LENARD, Berlin (1972).

[31] R.A. MINLOS and J.G. SINAI: cit. [23].

[32] G. GALLAVOTTI and S. MIRACLE: *Phys. Rev.*, **5**B, 2555 (1872); see also A. MARTIN–LÖF: *Comm. Math. Phys.*, **24**, 253 (1972).

[33] This theorem is due to R.A. MINLOS and J.G. SINAI: cit. [23]. Actually Minlos and Sinai prove a more difficult theorem under slightly different conditions. The adaptation of the deduction of the results given here can be found in G. GALLAVOTTI, A. MARTIN–LÖF: *Comm. Math. Phys.*, **25**, 87 (1972); or better in G. GALLAVOTTI, A. MARTIN-LÖF and S. MIRACLE–SOLÉ, cit. [30].

[34] G. GALLAVOTTI, A. MARTIN–LÖF: cit. [33].

[35] For a comparison of various old definitions of surface tension, new ones and a proof of their equivalence see D. ABRAHAM, G. GALLAVOTTI, A. MARTIN–LÖF: *Lettere al Nuovo Cimento*, **2**, 143 (1971).

[36] The results of this section are due to G. GALLAVOTTI and H. van BEYEREN, *Lettere al Nuovo Cimento*, **4**, 699 (1972), for the two dimensional case[17] and to R.

[16] Further applications came in [BGJS]. Duality has found many more applications, for instance a very recent one is in [BC].

[17] See also [Ga1].

DOBRUSHIN for the three dimensional case: to appear in *Comm. Math. Phys.*.[18]

[37] R.J. BURFORD and M.E. FISHER: *Phys. Rev.*, **156**, 583 (1967).

[38] The reader may consult the book by D. RUELLE, cit. [10], p. 125; R.A. MINLOS and J.G. SINAI, cit [23]; F.A. BEREZIN and J.G. SINAI, cit. [24]; R.L. DOBRUSHIN, *Funct. Anal. and Appl.*, **3**, 22 (1968); and the review article by J. GINIBRE: in *Colloques du CNRS*, (Gif-sur-Yvette, 1970).

[39] D. RUELLE: cit. [10], p.112.

[40] R.A. MINLOS and J.G. SINAI: *Teor. Mat. Fiz.*, **2**, No. 2, (1970); W.J. CAMP and M.E. FISHER: *Phys. Rev. Lett.*, **26**, 73 and 565 (1971).

[41] D. ABRAHAM: *Studies Appl. Math*, **50**, 71 (1971).

[42] O. LANFORD, D. RUELLE: cit. [13].

[43] Metastability should be a dynamical (*i.e.* non equilibrium) phenomenon as suggested by O. LANFORD, D. RUELLE: cit. [13]. An interesting rigorous treatment of the metastability phenomenon in the case of very weak and very long ranged forces can be found in J.L. LEBOWITZ and O. PENROSE: to appear.

[44] see the review paper by J. GINIBRE, cit. [38]. See also J.L. LEBOWITZ and G. GALLAVOTTI: *J. Math. Phys.*, **12**, 1129 (1971).

[45] D. RUELLE: *Phys. Rev. Lett.*: **27**, 1041 (1971).

[46] Some interesting results are to be expected in these cases. See for instance D. MERMIN: *Phys. Rev. Lett.*, **26**, 168 (1971).

[47] D. ABRAHAM: cit. [41].

[48] R.A. MINLOS and J.G. SINAI: cit. [40]; W.J. CAMP and M.E. FISHER: cit [40].

[49] W.J. CAMP and M.E. FISHER: cit [40].

GENERAL REFERENCES

D. RUELLE: *Statistical Mechanics* (New York (1969).

R.A. MINLOS: *Lectures in Statistical Mechanics*, in *Russian Math. Surveys*, **23**, 137 (1968).

The interested reader will be helped by the very recent and complete review of rigorous results (many of which refer to the Ising model and its generalizations) in the article:

R.B. GRIFFITHS: *Rigorous results and theorems*, to appear in *Phase transitions and critical points*, edited by. C. DOMB and M.S. GREEN (New York).

ADDITIONAL REFERENCES

[AM] Abraham, D.B., Martin-Löf, A.: *The transfer matrix for a pure*

[18] A nice and very simple proof of the rigidity in three dimensions was discovered by van Beyeren, [Be].

phase in the two dimensional Ising model, Communications in Mathematical Physics, **31**, 245, 1973.

[AR] Abraham, D.B., Reed, P.: *Interface profile of the Ising interface in two dimensions*, Communications in Mathematical Physics, **49**, 35, 1976.

[Ai] Aizenman, M.: *Translation invariance and instability of phase coexistence in the two dimensional Ising system*, Communications in Mathematical Physics, **73**, 83–94, 1980.

[Be] van Beyeren, H.: *Interface sharpness in the Ising system*, Communications in Mathematical Physics, **40**, 1–6, 1975.

[BC] Borgs, C., Chayez, J. T.: *The covariance matrix of the Potts model: A random cluster analysis*, UCLA preprint, September 1994.

[BGJS] Benettin, G., Gallavotti, G., Jona–Lasinio, G., Stella, A.: *On the Onsager–Yang value of the spontaneous magnetization*, Communications in Mathematical Physics: **30**, 45-54, 1973,

[BLPO] Bricmont, J., Lebowitz, J.L., Pfister, C., Olivieri, E.: Communications in Math. Phys., **66**, 1, 1979; and the first three authors in Communications in Math. Phys., **66**, 21, 1979 and **69**, 267, 1979.

[CCO] Capocaccia, D., Cassandro, M., Olivieri, E.: *A study of metastability in the Ising model*, Communications in Mathematical Physics, **39**, 185-205, 1974.

[CDR] de Conink J., Dunlop, F., Rivasseau, V.: *On the micoscopic validity of the Wulff construction and of the generalized Young equation*, Communincations in Mathematical Physics, **121**, 401, 1989.

[DKS] Dobrushin, R.L., Kotecky, R., Schlosman, S.: *A microscopic justification of the Wulff construction.*, Journal of Statistical Physics, **72**, 1, 1993. And *Wulff construction: a global shape from local interactions.*, American Mathematcial Society, 1992.

[Ga1] Gallavotti, G.: *Phase separation line in the two–dimensional Ising model*: Communications in Mathematical Physics, 27, 103–136, 1972.

[Ga2] Gallavotti, G. *Aspetti della Teoria Ergodica Qualitativa e Statistica del Moto*, Quaderno U.M.I., no. **21**, ed. Pitagora, Bologna, 1982, p.1–300.

[I] Isakov, S.N.: *Non analytic features of the first order phase transition in the Ising model*, Communications in Mathematical Physics, **95**, 427–443, 1984.

[Hi] Higuchi, Y.: *On the absence of non tramslationally invariant Gibbs states for the two dimensional Ising system*, in "Random fields", editors J. Fritz, J.L. Lebowitz and D. Szaz, North Holland, 1981.

[KM] Kotecky, R., Miracle-Solé, S.: *Roughening transition for the Ising model on a bcc lattice. A case in the theory of ground states*, Journal of Statistical Physics, **47**, 773, 1987.

[KO] Kotecky, R., Olivieri, E.: *Droplet dynamics for asymmetric Ising model*, J. Stat. Phys., **70**, 1121–1148, 1993.

[M] Miracle-Solè, S.: *On the microscopic theory of phase coexistence*, in "XIII Sitges Conference", Springer Lecture Notes in Physics, 1995.

[MOS] Martinelli, F., Olivieri, E, Scoppola, E.: *Metastability and exponential approach to equilibrium for low temperature stochastic Ising models*, J. Stat. Physics, **61**, 1105–1119, 1990.

[Pf] Pfister, C.: *Large deviations and phase separation in the two-dimensional Ising model*, Helvetica Physica Acta, **64**, 953– , 1991.

[PS] Pirogov, S.A., Sinai, J.G.: *Phase diagrams of classical lattice systems*, Theor. Math. Phys., **25**, 1185–1192, 1976; and **26**, 39–49, 1976.

[Ru] Russo, L.: *The infinite cluster method in the two dimensional Ising model*, Communications in Mathematical Physics, **67**, 251, 1979.

[Sc] Schonmann, R.H.: *The pattern of escape from metastability of a stochastic Ising model*, Communications in Mathematical Physics, **147**, 231–240, 1992.

[Wa] van der Waerden, B.L.: *The long range of the regular atom ordering in solid solutions*, (german), Zeitschrift für Physik, **118**, 473–488, 1941.

æ

RIGOROUS THEORY OF THE BOLTZMANN EQUATION IN THE LORENTZ GAS*

Giovanni Gallavotti**

Istituto di Fisica, Università di Roma
C.N.R., Gruppo Nazionale Analisi Funzionale

Abstract: The Boltzmann limit conjecture of Grad is discussed in general and proved for the Lorentz gas case (where the Boltzmann equation is linear). This is a reprint of an unpublished preprint of 1972, with one footnote added, one postscript (to quote the Lanford theorem), and improved with language editing. I reprint it in this form to make it accessible, as it has been quoted by other authors in later papers. The original preprint was commissioned for a book that eventually was not published.

1. — Introduction

The Boltzmann equation is an approximation to the "true" evolution equation: this is due to the fact that in its derivation the following assumptions are made [1,2]:
1) only binary collisions are considered
2) "Molecular chaos" is assumed at all times: i.e. the high order correlation functions can be expressed in terms of the one–particle distribution as:

$$f(\underline{r}_1\underline{v}_1, \underline{r}_2\underline{v}_2, \ldots, \underline{r}_n\underline{v}_n; t) = \prod_{i=1}^{n} f(\underline{r}_1\underline{v}_1;\ t) \tag{1.1}$$

* The first version appeared as a preprint: Nota Interna n. 358, Istituto di Fisica, Università di Roma, 10 feb 1972. The reprint has been deposited in the archive *mp_arc@math. utexas.edu*, # 93–304.

** Permanent address: Dipartimento di Fisica, Università di Roma.

3) in the computation of the collision term one disregards the fact that the molecules have a non vanishing dimension.

Having realized that the Boltzmann equation is only an approximation it becomes of interest to investigate if there are, at least, limiting cases in which it holds rigorously.

If n denotes the particle density and a the radius of interaction (i.e. a parameter proportional to the interaction range or to the square root of the scattering cross–section) then a critical examination of assumptions 1), 2), 3) suggests that the Boltzmann equation should hold rigorously in the limiting case $n \to \infty$, $a \to 0$ in such a way that [3]:

$$\text{i)} \quad na^2 \neq 0 \tag{1.2}$$

$$\text{ii)} \quad na^3 \to 0 \tag{1.3}$$

$$\text{iii)} \quad \text{“Molecular chaos” (1.1) is assumed at } t = 0 \tag{1.4}$$

In fact (1.2) says that the mean free path (m.f.p.) is finite (i.e. there are collisions); on the other hand (1.3) insures that, in the molecular scale, the gas is infinitely dilute (i.e. no particle can be found in a region of dimension a; hence the probability that a given particle collides with any other fixed particle vanishes *a priori* (although the particle will certainly suffer collisions because of (1.2)). The last facts prevent multiple collisions and the building up of correlations capable of destroying the molecularly chaotic character of the initial state.

To discuss rigorous results we need, however, a more precise statement of the conjecture that in the above limiting case the Boltzmann equation is rigorously true. This is done in the following lines.

Consider a gas of particles described, at $t = 0$, by a molecularly chaotic state (i.e. by a state such that the n–point correlation function factorizes as in (1.1)) with a one particle distribution:

$$f(\underline{r}, \underline{v};\ 0) = \frac{1}{a^2} f_0(\underline{r}, \underline{v}) \tag{1.5}$$

where $f_0(\underline{r}, \underline{v})$ is a given (a–independent) function.

Suppose that the gas of particles just introduced evolves through the action of a pair potential $\varphi_a(\underline{r})$ having the form:

$$\varphi_a(\underline{r}) = \varphi\left(\frac{|\underline{r}|}{a}\right) \tag{1.6}$$

where φ is a short range force (without hard core to avoid inessential complications in the notations and definitions) so that the differential scattering cross section in the solid angle Ω is of the form:

$$\sigma_a(\Omega) = a^2 \sigma(\Omega) \tag{1.7}$$

Let $f(\underline{r}_1, \underline{v}_1, \underline{r}_2, \underline{v}_2, \ldots, \underline{r}_n, \underline{v}_n;\ t)$ be the m–particle correlation function describing the state into which the initial state evolves in time t.

This function is, of course, no longer "chaotic" (i.e. of the form (1.1)). The "Boltzmann limit conjecture" (BLC) can be now formulated as:

Conjecture BLC:
For all fixed $t > 0$ and under "mild assumption" on $f_0(\underline{r}, \underline{v})$, the following limit exists:

$$\tilde{f}(\underline{r}_1\underline{v}_1, \underline{r}_2\underline{v}_2, \ldots, \underline{r}_m\underline{v}_m;\ t) = \lim_{a\to 0} a^{2m} f(\underline{r}_1\underline{v}_1, \ldots, \underline{r}_m\underline{v}_m;\ t) \tag{1.8}$$

and is "chaotic":

$$\tilde{f}(\underline{r}_1\underline{v}_1, \underline{r}_2\underline{v}_2, \ldots, \underline{r}_n\underline{v}_n;\ t) = \prod_{i=1}^{n} \tilde{f}(\underline{r}_i\underline{v}_i\ ;\ t) \tag{1.9}$$

and, further, $\tilde{f}(\underline{r}, \underline{v};\ t)$ verifies the Boltzmann equation:

$$\frac{\partial \tilde{f}}{\partial r}(\underline{r}, \underline{v}, t) + \underline{v} \cdot \frac{\partial \tilde{f}}{\partial r}(\underline{r}, \underline{v}; t) = \int d\,\underline{v}_1 \int d\,\underline{\omega} \cdot$$
$$\cdot\, |\underline{v} - \underline{v}_1| \sigma(\underline{\omega})(\tilde{f}(\underline{r}, \underline{v}'; t)\tilde{f}(\underline{r}_1, \underline{v}_1'; t) - \tilde{f}(\underline{r}, \underline{v}, t)\tilde{f}(\underline{r}, \underline{v}_1, t)) \tag{1.10}$$

with initial condition:

$$\tilde{f}(\underline{r}, \underline{v}, 0) = f_0(\underline{r}, \underline{v}) \tag{1.11}$$

where $\underline{v}', \underline{v}_1'$ are functions of $\underline{v}$, $\underline{v}_1$, $\underline{\omega}$ in such a way to conserve kinetic energy, linear momentum and to have a relative direction parallel to $\underline{\omega}$.

We refrain to state some examples of "mild assumptions" on $f_0(\underline{r}, \underline{v})$ since they should become clear in the course of a hypothetic proof of BLC.

We observe that, for the time being, it seems quite difficult to attack the problem of costructing a rigorous proof of the BLC. This is mainly due to the fact that the existence and stability theorems necessary for a proper mathematical definition of time evolutions of large systems are still far from being proved.

We stress that this is not a "technical point" but reflects our lack of understanding of some basic physical properties of the time evolution of large assemblies of particles (for an example of such problems see [5]; for an example of their applications see [6] and [7]).

In this paper we investigate the BLC in the case of simple models introduced by Lorentz [8] and used, for instance, to study the diffusion between gases of very different molecular weight [9] or, in its quantum version, to study the properties of a degenerate gas [10].

The Lorentz models are described in the next section and are such that the mathematical problems concerning existence and stability of the solutions of the microscopic equations of motion are very easily dealt with.

This mathematical simplicity, reflected also in the fact that the Boltzmann equation turns out to be linear, will enable us to push to the end the proof of the BLC at least in some cases.

2. — The Lorentz models

In the models there are two types of particles: the W–particles (wind–particles) and the T–particles (tree–particles).

The W–particles move through the space interacting only with the T–particles which, however, are supposed to be infinitely heavy compared to the W–particle and are supposed at rest and randomly distributed in space.

Each model is completely described by the $W - T$ interaction and by the T–particle distribution.

From now on we shall focus our interest to the case in which the T–particles are distributed as the space distribution of a perfect gas (Poisson distribution) with density n. We shall also assume that the T–particles are, with respect to the W–particles, hard spheres of radius a, reflecting the W–particles on their surface.

The assumed tree distribution is such that the probability for finding inside a given region Λ, with volume $V(\Lambda)$ exactly N tree particles, and for finding them in the infinitesimal cubes $d\underline{c}_1, \dots, d\underline{c}_N$ around $\underline{c}_1, \dots, \underline{c}_N$, is:

$$f_\Lambda(\underline{c}_1, \dots, \underline{c}_N) d\underline{c}_1, \dots, d\underline{c}_N = e^{-nV(\Lambda)} \frac{n^N}{N!} d\underline{c}_1, \dots, d\underline{c}_N \tag{2.1}$$

Note that, since the T–particles are hard spheres only with respect to the W–particles but not with respect to the each other, there are configurations $\underline{c}_1, \dots, \underline{c}_N$ of trees in which the hard spheres overlap, (for some comments on this point see §6).

If $x = (\underline{p}, \underline{q})$ is the W–particles phase space coordinate ($\underline{p}$= velocity, $\underline{q}$= position) the symbol:

$$S_t^{\underline{c}_1, \dots \underline{c}_N} x \tag{2.2}$$

will denote the W–particle $x' = (\underline{p}', \underline{q}')$ into which x evolves in time t in the presence of N tree–particles located at $\underline{c}_1, \dots, \underline{c}_N$.

The symbol $\omega(\underline{p})$ will denote the direction of $\underline{p}$. The symbol $\hat{x}$ will denote the pair $(\omega(\underline{p}), q)$ if $x = (\underline{p}, \underline{q})$.

Since the velocity $|\underline{p}|$ is conserved it is clear that $S_t^{\underline{c}_1, \dots, \underline{c}_N} x$ depends only on the trees located within a distance $(|\underline{p}|t + a)$ from $\underline{q}$. The symbols:

$$\left(S_t^{\underline{c}_1, \dots, \underline{c}_N} x\right)_1 \ , \ \left(S_t^{\underline{c}_1, \dots, \underline{c}_N} x\right)_2 \ , \ \omega\left(\left(S_t^{\underline{c}_1, \dots, \underline{c}_N} x\right)_1\right) \tag{2.3}$$

will, respectively, denote the velocity, position and momentum direction of (2.2). The following symbols will occur frequently:

$$S_t^{\underline{c}_1,\dots,\underline{c}_N}\hat{x} = \left(\omega\left(\left(S_t^{\underline{c}_1,\dots,\underline{c}_N}x\right)_1\right)\ ,\ \left(S_t^{\underline{c}_1,\dots,\underline{c}_N}x_m\right)\right) \tag{2.4}$$

Similarly we can give a natural meaning to the evolution of m W–particles:

$$S_t^{\underline{c}_1,\dots,\underline{c}_N}(x_1,\dots,x_m) = \left(S_t^{\underline{c}_1,\dots,\underline{c}_N}x_1,\dots,S_t^{\underline{c}_1,\dots,\underline{c}_N}x_m\right) \tag{2.5}$$

which takes into account the fact that there are no $W-W$ interactions.

3. — The Boltzmann limit for the Lorentz gas

It is easy to derive the Boltzmann equation for W–particles in the case of the above described Lorentz gas. It is not difficult to realize that the assumptions to be made in order to derive the Boltzmann equation are essentially the same as conditions 1), 2), 3), of section 1. They are:
i) a W–particle never hits twice the same particle;
ii) molecular chaos is assumed;
iii) the size of the T–particles is negligible.

Here by "chaotic" W–particle state we again mean a state such that the W–particle correlation functions are a product of one W–particle distribution which are independent on the T–particle distribution: more precisely a chaotic state is such that the probability distribution for finding a certain configuration C of T–particles and a set of W–particles in $x_1,\dots,x_m$ has the form: $p(C)\prod_{i=1}^m f(x_i)$, where $p(C)$ denotes the Poisson distribution (2.1) and this is interpreted as 0 if any wind particle is inside the hard cores of C.*

Clearly assumption i), ii), iii) can be only approximately true.

Let us formulate the BLC for the Lorentz gas. Assume that the initial W–particles state has the form:

$$f(x_1,\dots,x_m;\ 0) = \int_{C\ \mathrm{comp}\,(x_1,\dots,x_m)} p(C)\prod_{i=1}^m f_0(x_i) \tag{3.1}$$

* More explicitly this means the following. Let p be the probability of finding the W particles in a infinitesimal cube $dx_1\dots dx_m$ around the configuration $X=(x_1,\dots,x_m)$ in the box Λ_0, and a tree configuration in the infinitesimal cube $d\underline{c}_1\dots d\underline{c}_M$ around $C=(\underline{c}_1,\dots,\underline{c}_M)$ in the box Λ, assuming it wider by an amount a than Λ_0, at least. Here $x_i=(\underline{p}_i,\underline{q}_i)$. Then p is the product of (2.1) times $M!^{-1}(\prod_{i=1}^M f_0(x_i)\,dx_i)\,e^{-\int_C f_0(\underline{\xi})d\underline{\xi}}$, where $\underline{\xi}=(\underline{p},\underline{q})$ and $\int_C d\underline{\xi}$ means integration over $\underline{p}$ and over the $\underline{q}\in\Lambda_0$ which are outside the hard spheres centered on $C=(\underline{c}_1,\dots,\underline{c}_N)$. In other words the W particles also have a Poisson distribution, in the region outside the T particles, with a density function f_0.

where $f_0(x)$ is a given function of x and the "integral" is the "sum" over all the T–particle configurations compatible with $x_1, \dots, x_m$ (i.e. over the C's such that no W–particle is located inside the hard core of a T–particle).

Note that (3.1) is not a product state for the W–particles: this difference with respect to section 1 arises because here we have hard core interactions (which, for simplicity, were not considered in section 1).

Consider the state obtained by evolving the initial state (3.1):

$$f(x_1, \dots, x_m;\ t) = \int_{C\ \text{comp}\ (x_1,\dots,x_m)} p(C) \prod_{i=0}^{m} f_0(S^C_{-t} x_j) \tag{3.2}$$

and then let the T–particle density n tend to infinity and the hard core $W-T$ radius tend to zero in such a way that $na^3 \to 0$ but $na^2 \to l \neq 0, \infty$.

Imagining that the solid angle integration is normalized to $\int d\underline{\omega} = 1$, the BLC becomes:

Conjecture BLC:

If $t \geq 0$ and under "mild assumptions" on f_0, the following limit exists:

$$\lim_{\substack{na^3 \to 0 \\ na^2 \to\ \text{const}\ \neq 0,\infty}} f(x_1, \dots, x_m;\ t) = \tilde{f}(x_1, \dots, x_m;\ t) \tag{3.3}$$

and:

$$\tilde{f}(x_1, \dots, x_m;\ t) = \prod_{i=1}^{m} \tilde{f}(x_i;\ t) \tag{3.4}$$

and $\tilde{f}(x;t)$ verifies the Boltzmann equation:

$$\frac{\partial \tilde{f}}{\partial t}(x,t) + \underline{p} \cdot \frac{\partial \tilde{f}}{\partial \underline{q}}(x,t) = \lambda^{-1} |\underline{p}| \int (\tilde{f}(x',t) - \tilde{f}(x,t)\sigma(\underline{\omega}))\, d\underline{\omega} \tag{3.5}$$

where $x = (\underline{p}, \underline{q}), x' = (\underline{p}';\ \underline{q})$ and $\underline{p}'$ is a vector with the same length as $\underline{p}$ but forming with it an angle $\underline{\omega}$; $a^2\sigma(\underline{\omega}) = a^2$ is the scattering cross section of a hard sphere with radius a and $\lambda^{-1} = 4\pi n a^2$.

A similar conjecture can be formulated in a two–dimensional model; here the solid angle $\underline{\omega}$ has to be replaced by the deflection angle β (see Fig.(5.8) below) and $\sigma(\underline{\omega})$ by $\sigma(\psi) = \frac{\pi}{2} \sin \frac{\beta}{2}$ and $\lambda^{-1} = 2an$. Of course the Boltzmann limit will be, in this case, $na^2 \to 0$, $2na \to \lambda^{-1} \neq 0, \infty$.

In the next sections we construct a proof of the above conjecture in the two–dimensional case. The three–dimensional case could be treated along the same lines as it will become apparent from the proofs.

4. — Results on the BLC

Assume the spatial dimension to be two. The direction $\omega(\underline{p})$ will be in this case the angle ϑ between $\underline{p}$ and a fixed axis.

The function $f_0(x)$ will be thought as $f_0(|\underline{p}|, \omega(\underline{p}), \underline{q})$, if $x = (\underline{p}, \underline{q})$, and we can write:

$$f_0(|\underline{p}|, \omega(\underline{p}), \underline{q}) = \int d\,\underline{q}'\, d\omega' f_0(|\underline{p}|, \omega', \underline{q}')\delta(\underline{q} - \underline{q}')\delta(\omega(\underline{p}) - \omega') \tag{4.1}$$

we shall shorten (ω', q') as ξ, $d\underline{q}'d\omega'$ as $d\xi$, $\delta(\underline{q} - q')\delta(\omega(\underline{p}) - \omega')$ as $\delta(x - \xi)$. Hence, by using definition (2.4), the (3.2) becomes, for $m = 1$:

$$f(x;t) = \int d\,\xi f_0(|\underline{p}|, \xi) \int_{C\ \mathrm{comp}\, x} \delta(S^C_{-t}\hat{x} - \xi)p(C) \tag{4.2}$$

It is therefore useful to consider the Green's function:

$$g(\xi; x; t) = e^{\pi n a^2} \int_{C\ \mathrm{comp}\, x} p(C)\delta(S^C_{-t}\hat{x} - \xi) \tag{4.3}$$

where the factor $e^{\pi n a^2}$ has been introduced for normalization purposes (note that it tends to 1, in the Boltzmann limit).

It is easily checked that:

$$\begin{aligned} g(\xi; x; 0) =& \delta(\hat{x} - \xi), \quad \int g(\xi; x; t)d\xi \equiv 1 \\ f(x,t) =& e^{-na^2} \int d\,\xi f_0(|\underline{p}|, \xi)g(\xi; x; t) \end{aligned} \tag{4.4}$$

we shall show that as $na^2 \to 0$, $2na \to \lambda^{-1} \neq 0, \infty$ the function $g(\xi; x; t)$ will tend to a limit $\tilde{g}(\xi; x; t)$ which verifies the two dimensional analogue of equation (3.5) with initial condition $\tilde{g}(\xi; x; 0) = \delta(\hat{x} - \xi)$ and $|\underline{p}|$ fixed.

The linearity of (3.5), and of the third (4.4), will imply, under suitable assumptions on f_0, that also $\tilde{f}(x,t)$ verifies (3.5). We will not insist in discussing in which sense $g(\xi; x; t)$ converges to $\tilde{g}(\xi; x; t)$. It will appear from the proofs below that at least $g(\xi; x; t)$ converges to $\tilde{g}(\xi; x; t)$ pointwise for $t \neq 0$, and in the sense of the distributions for all $t \geq 0$. However a close examination of the proof will provide evidence against any uniformity of the convergence in t, unless t is restricted to a bounded interval (for further remarks on this point see section 6).

Under the above convergence conditions the "mild assumptions" in BLC could, for instance, be the continuity and boundedness of f_0.

5. — Proof

The proof is based on several simple changes of variables in (4.3).

Let $x = (\underline{p}, \underline{q})$ and let $R(x,t)$ be the sphere with center $\underline{q}$ and radius $(|p|t + a)$; then $S^c_{-t}x$ depends only on the T–particles in c contained in $R(x,t)$. Hence the integral (4.3) can be explicitly written as:

$$g(\xi;x;t) = \\ = e^{\pi n a^2} \sum_{M=0}^{\infty} \int_{R(x,t)^M} e^{-nV(R(x,t))} \frac{n^M}{M!} \delta\left(S^{\underline{c}_1,\dots,\underline{c}_M}_{-t}\hat{x} - \xi\right) d\underline{c}_1 \dots d\underline{c}_M \tag{5.1}$$

where $V(R(x,t))$= area of $R(x,t)$ and where use has been made of the assumed Poisson distribution of the T–particles (2.1).

Note that, in general, not all the T–particles $\underline{c}_1, \dots, \underline{c}_M$ in (5.1) will be hit by the trajectory $S^{\underline{c}_1,\dots,\underline{c}_M}_{-t}x$ $0 \le \tau \le t$. Let $A_{x,t,N}$ denote the set of configurations $\underline{c}_1, \dots, \underline{c}_N$ of N T–particles such that a W–particle with initial coordinate x hits, in the time t, all the N particles in $\underline{c}_1, \dots \underline{c}_N$ at least once. We deduce from (5.1), see the figure below:

(5.2)

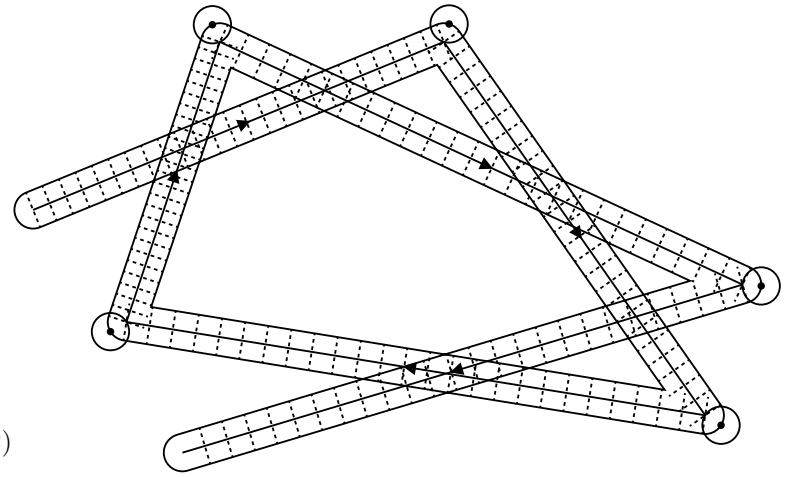


Fig. (5.2): The set $P(t; \underline{c}_1, \dots, \underline{c}_N)$ is the dashed region. The circles represent trees $\underline{c}_1, \dots, \underline{c}_N$, $(N = 5)$ and the length of the trajectory in the dashed region is $|\underline{p}|t$.

$$g(\xi;x;t) = e^{\pi n a^2} \sum_{N=0}^{\infty} \int_{A_{x,t,N}} n^N \frac{d\underline{c}_1,\ldots,d\underline{c}_N}{N!} \chi_{\underline{c}_1,\ldots,\underline{c}_N}(x) \cdot \delta\left(S_{-t}^{\underline{c}_1,\ldots\underline{c}_N}\hat{x} - \xi\right) \cdot$$
$$\cdot \left[\sum_{M=N}^{\infty} \int_{\mathcal{R}} \frac{d\underline{c}'_1,\ldots,d\underline{c}'_{M-N}}{(M-N)!}\, n^{M-N} e^{-nV(R(x,t))}\right]$$
$$\mathcal{R} \equiv R(x,t)^{M-N};\ \underline{c}'_1,\ldots,\underline{c}'_{M-N} \in P(t;\underline{c}_1,\ldots,\underline{c}_N) \tag{5.3}$$

where $\chi_{\underline{c}_1,\ldots,\underline{c}_N}(x)$ is 1 if x is compatible with the hard cores of $\underline{c}_1,\ldots,\underline{c}_N$ and 0 otherwise: the region $P(t;\underline{c}_1,\ldots,\underline{c}_N)$ is the tube like region, (see fig. (5.2)), swept by an ideal T–particle when its center is moved along the path $S_{-\tau}^{\underline{c}_1,\ldots,\underline{c}_N}x$, $0 \le \tau \le t$.
The sum within square brackets in (5.2) can be performed (since the integrals are trivials) and yields:

$$e^{-nV(P(t;\underline{c}_1,\ldots,\underline{c}_N))} \tag{5.4}$$

so that $g(\xi,x;t)$ is:

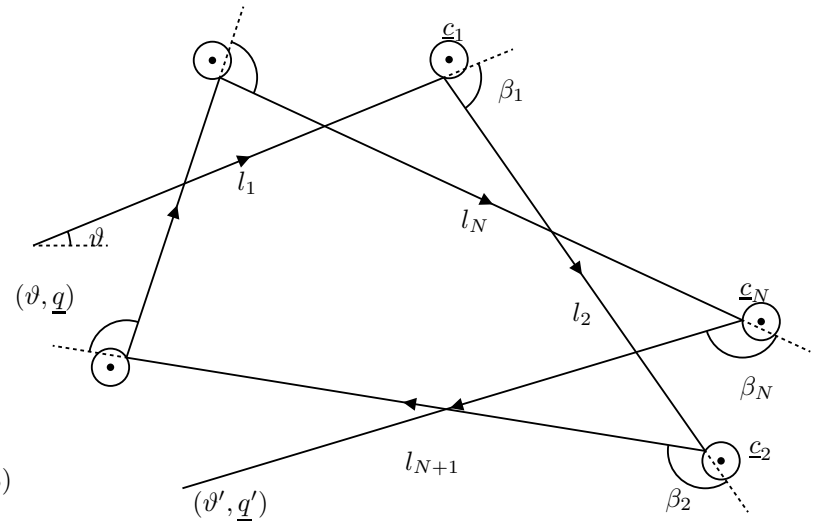
$$g(\xi;x;t) = e^{n\pi a^2} \sum_{N=0}^{\infty} \int_{A_{x,t,N}} n^N e^{-nV(P(t;\underline{c}_1,\ldots,\underline{c}_N))} \frac{d\underline{c}_1\ldots d\underline{c}_N}{N!} \cdot \delta\left(S_{-t}^{\underline{c}_1,\ldots,\underline{c}_N}\hat{x} - \xi\right) \tag{5.5}$$

The reader should remark the very simple probabilistic meaning of this equation which makes it almost self–evident [11]: the T–particles in $A^1_{x,t,N}$ can be hit more than once in the time t. Divide $A_{x,t,N}$ as $A^1_{x,t,N} \cup A'_{x,t,N}$ where $A^1_{x,t,N}$ is the set of T–configurations in $A_{x,t,N}$ such that all their T–particles are hit just once by the trajectory $S_{-\tau}^{\underline{c}_1,\ldots,\underline{c}_N}x$ $0 \le \tau \le t$.
To this decomposition of $A^1_{x,t,N}$ corresponds a decomposition $g(\xi;x;t) = g_1(\xi;x;t) + g'(\xi;x;t)$ with:

$$g_1(\xi;x;t) = e^{\pi n a^2} \sum_{N=0}^{\infty} \int_{A^1_{x,t,N}} n^N \frac{d\underline{c}_1,\ldots,d\underline{c}_N}{N!} \chi_{\underline{c}_1,\ldots,\underline{c}_N}(x) \cdot \delta\left(S_{-t}^{\underline{c}_1,\ldots,\underline{c}_N}\hat{x} - \xi\right) e^{-nV(P(t;\underline{c}_1,\ldots,\underline{c}_N))} \tag{5.6}$$

We now perform the change of variables, illustrated in fig. (5.8), from the $2N$ variables $\underline{c}_1,\ldots,\underline{c}_N$ to the new $2N+1$ variables $l_1,\ldots,l_{N+1}$, $\beta_1,\ldots,\beta_N$; we get:

$$\frac{d\underline{c}_1,\ldots,d\underline{c}_N}{N!} = a^N \delta\left(\sum_{i=1}^{N+1} l_i - |\underline{p}|t\right) \left(\prod_{i=1}^{N+1} dl_j\right) \prod_{j=1}^{N} \left(\frac{d\beta_j}{2}\, \sin\frac{\beta_j}{2}\right) \tag{5.7}$$



(5.8)

Hence the N^{th} order contribution to (5.6) is given by (if $x = (p, q) = (|p|, \omega(\underline{p}), \underline{q})$, $\xi = (|\underline{p}|, \vartheta', q')$:

$$* \; e^{\pi n a^2} (2na)^N \int_0^\infty \prod_{i=1}^{N+1} dl_i \int_0^{2\pi} \prod_{i=1}^{N} \left(\sin \frac{\beta_i}{2} \, \frac{d\beta_i}{4} \right) \delta \left(\sum_{i=1}^{N+1} l_i - |\underline{p}| t \right) \cdot$$
$$\cdot \, \delta \left(\sum_{i=1}^{N+1} \underline{l}_i - (\underline{q}' - \underline{q}) \right) \delta \left(\sum_{i=1}^{N} \beta_i - (\vartheta' - \omega(\underline{p})) \right) e^{-nV(P(t; \underline{c}_1, \ldots, \underline{c}_N))} \quad (5.9)$$

where $\underline{l}_i$ are the vectors represented by arrows in fig. (5.2) ($|\underline{l}_1| = l_i$); the * in (5.9) means that there is an extra condition on the integration region. It is the condition that none of the spheres of radius a around $\underline{c}_1, \ldots, \underline{c}_N$ has intersection with the straight segments of the broken line representing the trajectory in fig. (5.8) (i.e. this is the condition that $\underline{c}_1, \ldots, \underline{c}_N$ really belongs to $A^1_{x,t,N}$). Of course in (5.9), $\delta \left(\sum_{i\equiv 1}^{N} \beta_i - (\vartheta' - \omega(\underline{p})) \right)$ means $\sum_{h=-\infty}^{+\infty} \delta(\sum_i \beta_i - (\vartheta' - \omega(\underline{p})) - 2\pi h)$.

In the limit $na^2 \to 0$, $2na \to \lambda^{-1} \neq 0, \infty$ the restrictions indicated by the * in (5.9) become unimportant and $nV(P(t; \underline{c}_1, \ldots, \underline{c}_N))$ simplifies enormously:

$$nV(P(t; \underline{c}_1, \ldots, \underline{c}_N)) \to 2na \sum_{j=1}^{N+1} l_j = \lambda^{-1} |p| t \quad (5.10)$$

Hence the limit $\tilde{g}(\xi;x;t)$ as $na^2 \to 0$ and $2na \to \lambda^{-1} \neq 0, \infty$ of $g(\xi;x;t)$ is:

$$\sum_{N=0}^{\infty} \lambda^{-N} \int_0^{\infty} \int_0^{2\pi} \Big(\prod_{i=1}^{N} \sin \frac{\beta_j}{2} \frac{d\beta_j \, dl_j}{4} \Big) \delta(\sum l_j - |\underline{p}|t) \cdot \\ \cdot \delta \left(\sum_i \underline{l}_i - (\underline{q}' - \underline{q}) \right) \cdot \delta \left(\sum_i \beta_i - (\vartheta' - \omega(\underline{p})) \right) \cdot e^{-\lambda^{-1}|p|t} \tag{5.11}$$

@

In the derivation of equation (5.11) we have systematically disregarded convergence problems connected with the summation over N, M etc., since they are trivial as a consequence of the presence of the factorials and of the boundedness of the integration regions.

The limit (5.11) is pointwise for $t \neq 0$ and it could be checked that it holds also in the sense of the distributions for $t \geq 0$.

Furthermore it could be checked that for $t > 0$ the function $g(\xi, x, t) \geq g_1(\xi, x, t)$ is bounded above by a $L_1(d\xi)$ function; hence the limit (5.11) holds also in the $L_1(d\xi)$ sense. Finally, by direct computation, it follows from (5.11) that:

$$\int \tilde{g}(\xi;x;t) d\xi \equiv 1 \tag{5.12}$$

and this fact, together with the above convergence properties and (4.4), implies the validity of the limit relation: $\lim_{\substack{na^2 \to 0 \\ 2na \to \lambda^{-1}}} g(\xi;x;t) = \tilde{g}(\xi;x;t)$ in $L_1(d\xi)$ for $t > 0$; furthermore it could be proved that this limit holds, for $t \geq 0$, in the sense of the distributions.

It is known [12] that (5.11) is a solution of the Boltzmann equation (and this can be checked directly by substituting $\tilde{g}$ into (3.5)), with initial condition $\tilde{g}(\xi;x;0) = \delta(\hat{x} - \xi)$ and $|p|$ fixed.

To complete the proof of the BLC it remains to deal with the m–particle distributions. However we skip this point since it involves straightforward calculations based on changes of variable of the type illustrated in fig. (5.8).

6. — Concluding remarks

In the preceding sections we have described a proof of the Boltzmann limit conjecture in the case of a two–dimensional Lorentz gas with hard core $W - T$ interactions and free gas distribution of the T–particles. The generalization to three dimensions would be trivial.

A less trivial generalization would be obtained by keeping the hard core $W - T$ interaction but assuming that the T–particles are spatially distributed as if they were a gas of hard spheres with hard core size being

proportional to the the $W-T$ radius. Other generalizations are conceivable in the direction of allowing soft $W-T$ particle interactions and more general T–particle distributions. So far none of these generalizations have been attempted.

Much more difficult and interesting would be the treatment of the*Knudsen model*, in which the T–particles are allowed to move without suffering changes in their momentum in the collisions with the W–particles.

Had we done the calculations associated with the proof of equation (3.4), we would have also found evidence of a lack of uniformity of the Boltzmann limit in the number m of W–particles even at fixed t: the larger m is the closer one has to get near the Boltzmann limit in order to see factorization of the W–particle correlations.

We also wish to remark that even when the Boltzmann limit conjecture is true, one cannot expect that the function $\tilde{f}(\underline{r},\underline{v},t)/a^2$ (see section 1) is a good approximation to $f(\underline{r},\underline{v},t)$ for large t: in fact one intuitively expects that for times of the order of $t_{m.f.p.}/na^3$ some non trivial correlations will start building up thus destroying themolecular chaos and spoiling the validity of the Boltzmann equation.

This last remark is quite deceiving since it tells us that we cannot use, *without further assumptions*, the Boltzmann equation to investigate the long time behaviour and, in particular, to compute the transport coefficients. From a rigorous point of view we cannot even be sure that the lowest order in na of the transport coefficients is correctly given by the value obtained in the Boltzmann limit. However it seems reasonable that this is, indeed, the case at least if the dimension of the space is larger than two (in one–dimension a simple counter example can be found by using soluble models [4]; in this case, however, the Boltzmann equation is a priori expected not to be a good approximation).

For further readings on the Lorentz gas see ref. [13].

The idea of the Boltzmann limit is clearly stated in [3]; the present proof in the case of the Lorentz gas is done in ref. [12] (for the case of $\tilde{g}(\underline{\xi};x;t)$ only) and was inspired by discussions and suggestions from J.L. Lebowitz.

Postscript:

The Boltzmann limit conjecture has been proved a few years after this paper was written. It is due to O. Lanford, [14] (1974), and it holds under some very reasonable restrictions on the initial data and for a hard spheres system, but with a still standing limitation on the time interval of validity. The time interval is strictly positive, but it is a small fraction ε of the mean free flight time t_0 (which is defined as the ratio of the mean free path

*over a mean velocity computed in the initial state): $\varepsilon \sim \frac{1}{5}$. This clearly did put an end (or at least it should have) to the diatribes on whether the irreversibility can or cannot be deduced from a microscopically reversible mechanical model (*Boltzmann vindicatus est*); but it left open the question of the mathematical justification of the validity of the Boltzmann equation over the time scales on which it is usually applied, going beyond the mean free time by several orders of magnitude. There has been, since, one case in which the proof of the BLC has been pushed to infinite time, [15].*

— References

[1] S. Chapman, T. Cowling, *The Mathematical Theory of Non–Uniform Gases*, Cambridge University Press, 1953, p. 46.
[2] Cohen, E.G.D.: *The kinetic theory of dilute gases*, in "Transport Phenomena in Fluids", H. Hanley ed., Ch. VI, 119–156; Dekker, New York, 1969.
[3] H. Grad, *Principles of the kinetic theory of gases*, in *Handbuch der Physik*, vol. XII, p. 205–294, see p. 214, ed. S. Flügge, Springer–Verlag, 1958.
[4] J. Lebowitz, J. Percus, *Kinetic equations and exactly solvable one dimensional systems*, Physical Review, **155**, 122–138, 1966; see also J. Lebowitz, J. Percus, J. Sykes, *Time evolution of the total distribution function of a one dimensional systems of hard rods*, Physical Review, **171**, 224–235, 1968.
[5] O. Lanford, *The classical mechanics of one dimensional systems of infinitely many particles.I. An existence theorem*, Communications in Mathematical Physics, **9**, 176–191, 1968.
[6] O. Lanford, *The classical mechanics of one dimensional systems of infinitely many particles.II. Kinetic theory*, Communications in Mathematical Physics, **11**, 257–292, 1969.
[7] G. Gallavotti, O. Lanford, J. Lebowitz: *Thermodynamic limit of time–dependent correlation functions for one–dimensional systems*: Journal of Mathematical Physics: **11**, 2898–2905, 1970
[8] S. Chapman, T. Cowling, *loc. cit.*, p. 187.
[9] S. Chapman, T. Cowling, *loc. cit.*, p. 256.
[10] S. Chapman, T. Cowling, *loc. cit.*, p. 309.
[11] G. Gallavotti, *Time evolution problems in Classical Statistical Mechanics and the Wind–tree–model*: in "Cargese Lectures in Physics", vol. IV,

ed. D. Kastler, Gordon Breach, Paris, 1970, pp.257–275; see p. 271–272, formula (6.2).
[12] G. Gallavotti, *Divergences and approach to equilibrium in the Lorentz and the Wind–tree models*, Physical Review, **185**, 308–322, 1969.
[13] A. Weijland, J. Van Leeuwen, *Non analytic behaviour of the diffusion coefficient of a Lorentz gas*, Physica, **36**, 457–490, 1967; and **38**, 35–47, 1968; E. Hauge, E. Cohen, *Normal and abnormal diffusion in Ehrenfests's wind tree model*, Journal of Mathematical Physics, **8**, 397–414, 1969; W., Hogey, *Convergent generalizations of the Boltzmann equation for a hard sphere gas*, Physical Review, **185**, 210–218, 1969.
[14] O. Lanford, *Time evolution of large classical systems*, in "Dynamical systems, theory and applications", p. 1–111, ed. J. Moser, Lecture Notes in Physics, vol. 38, Springer Verlag, 1974.
[15] M. Pulvirenti, *Global validity of the Boltzmann equation for two and three dimensional rare gas in vacuum*, Communications in Mathematical Physics, **113**, 79–85, 1987.

Ergodicity, ensembles, irreversibility in Boltzmann and beyond.[!]

Versione originale dell' articolo apparso su Journal of Statistical Physics, **78***, 1571–1589, 1995; autore: Giovanni Gallavotti*

Abstract:*the implications of the original misunderstanding of the etymology of the word "ergodic" are discussed, and the contents of a not too well known paper by Boltzmann are critically examined. The connection with the modern theory of Ruelle is attempted.*

Ruelle §**1***The etymology of the word "ergodic" and the heat theorems.*

Trying to find the meaning of the word "ergodic" one is led to a 1884 paper by Boltzmann, [B84].[1] This paper by Boltzmann is seldom quoted [2] and no english translation is available yet. But I think that this is one of the most interesting papers of Boltzmann: it is a precursor of the work of Gibbs, [G], on the ensembles, containing it almost entirely (if one recalls that the equivalence of the canonical and microcanonical ensembles was

[!] expanded and revised version of a conference read at the celebration of the 150^{th}–anniversary of the birth of Boltzmann, Vienna, 24 february, 1994; this paper is archived in *mp_arc@math.utexas.edu*, #94-66

[1] see the footnote of S. Brush in his edition, [Bo2], of the Lectures on Gas Theory, on p. 297 (§32): here the Boltzmann's paper is quoted as the first place where the word is introduced, although the etymology is taken from the Erhenfests' paper, which is incorrect on this point: see [EE], note #93, p.89, (where also the first appearance of the word is incorrectly dated and quoted).

[2] I found only the Brush's reference in [1], and a partial account in [Br1], p.242 and p. 368, before my own etymological discussion, appeared in print in [G1] after several years of lectures on the subject. My discussion was repeated in [G2] and [G3]. More recently the paper has been appropriately quoted by [Pl], unaware of my analysis. The paper was discussed also by [Ma], see footnote [9] below.

already established (elsewhere) by Boltzmann himself, at least in the free case [B66],[B68]), and I will try to motivate such statement.

The paper stems from the fundamental, not too well known, work of Helmholtz, [He1], [He2], who noted that *monocyclic* systems[3] could be used to provide models of thermodynamics in a sense that Boltzmann undertakes to extend to a major generalization.

After an introduction, whose relative obscurity has been probably responsible for the little attention this paper has received, Boltzmann introduces the notion of "stationary" probability distribution on the phase space of N interacting particles enclosed in a vessel with volume V. He calls a family $\mathcal{E}$ of such probabilities a *monode*, generalizing an "analogous" concept on monocyclic systems.[4]

In fact the orbits of a monocyclic system can be regarded as endowed with a probability distribution giving an arc length a probability proportional to the time spent on it by the motion: hence their family forms a family of stationary probability distributions.

Etymologically this undoubtedly[5] means a family of stationary distributions with a "unique nature", (each consisting of systems with a "unique nature", differing only by the initial conditions), from μόνος and εἶδος, with a probable reference to Plato and Leibnitz. Plato and Leibnitz.

Then the following question is posed. Given an element μ of a monode $\mathcal{E}$, also called a monode by Boltzmann, we can compute the average values of various observables, *e.g.* average kinetic energy, average total energy, average momentum transfer per unit time and unit surface in the collisions with the vessel walls, average volume occupied and density, denoted,

[3] this is what we call today a system whose phase space contains only periodic orbits, or cycles: *i.e.* essentially a one dimensional conservative system.

[4] in fact Boltzmann first calls a monode just a single stationary distribution regarded as an ensemble. But sometimes later he implicitly, or explicitly, thinks of a monode as a collection of stationary distributions parameterized by some parameters: the distinction is always very clear from the context. Therefore, for simplicity, I take here the liberty of calling "monode" a collection of stationary distributions, and the individual elements of the collection will be called "elements of the monode". The etymology that follows, however, is more appropriate for the elements of the monodes, as they are thought as consisting of many copies of the same system in different configurations. By reading the Boltzmann's analysis one can get the impression, see p. 132 of [B84], that the word monode had been already introduced by Maxwell, in [M]: however the reference to Maxwell is probably meant to refer to the notion of stationarity rather than to the word monode which does not seem to appear in [M].

[5] of course one can doubt (on this as well as on many other things).

respectively:

$$T = \frac{1}{N}\langle K\rangle_\mu, \quad U = \langle K + \Phi\rangle_\mu, \quad p, \quad V, \quad \rho = \frac{N}{V} \tag{1.1}$$

where Φ denotes the potential interaction energy and K the total kinetic energy. We then imagine to vary μ in the monode $\mathcal{E}$, by an infinitesimal amount (this meanss changing any of the parameters which determine the element). **Question:** *is it true that the corresponding variations dU and dV are such that*:

$$\frac{dU + p\,dV}{T} \quad \text{is an exact differential} \quad dS\ ? \tag{1.2}$$

In other words is it true that the above quantities, defined in purely mechanical terms, verify the same relation that would hold between them if, for some thermodynamic system, they were the thermodynamic quantities bearing the same name, with the further identification of the average kinetic energy with the absolute temperature?[6] If so the monode would provide a "mechanical model of thermodynamics" extending, by far, the early examples of Helmholtz on monocyclic systems.

Thus Boltzmann is led to the following definition:

Definition: *a monode $\mathcal{E}$ is called an* orthode *if the property described by (1.2) holds.*

Undoubtedly the etymology of "orthode" is ὀρθός and εἶδος, *i.e.* "right nature".

I find it almost unbelievable that such a deep definition has not been taken up by the subsequent literature. This is more so as Boltzmann, in the same paper, proceeds to discuss "examples" of mechanical models of thermodynamics, *i.e.* examples of orthodic monodes.

It has, certainly, not escaped the reader that an orthodic monode (or orthode) is what we call today an *equilibrium ensemble*. And the above orthodicity concept is still attributed to Gibbs, see [Br1], p. 242).

[6] that the temperature should be identified with the average kinetic energy per particle was quite well established (for free gases) since the paper by Clausius, [C], and the paper on the equipartition of kinetic energy by Boltzmann, [B68] (in the interacting cases); see the discussion of it in Maxwell's last scientific work, [M]. The latter paper is also very interesting as Maxwell asks there whether there are other stationary distributions on the energy surface, and tries to answer the question by putting forward the ergodic hypothesis.

The examples of orthodes discussed by Boltzmann in his paper are the *holode* and the *ergode* which are two ensembles whose elements are parameterized with two parameters β, N or U, N, respectively. Their elements are:

$$\mu_{\beta,N}(d\underline{p}d\underline{q}) = \frac{d\underline{p}_1 \dots d\underline{p}_n d\underline{q}_1 \dots d\underline{q}_n}{const} e^{-\beta(K+\Phi)} \tag{1.3}$$

and:

$$\mu_{U,N}(d\underline{p}d\underline{q}) = \frac{d\underline{p}_1 \dots d\underline{p}_n d\underline{q}_1 \dots d\underline{q}_n}{const} \delta(K(\underline{p}) + \Phi(\underline{q}) - U) \tag{1.4}$$

Boltzmann proves that the above two ensembles are both orthodes! thus establishing that the canonical and the microcanonical ensembles (using our modern terminology) are equilibrium ensembles and provide mechanical models of thermodynamics.[7]

Boltzmann's proof makes use of the auxiliary (with respect to the above definition) notion of heat transfer: in the canonical case it yields exactly the desired result; in the microcanonical it is also very simple but somehow based on a different notion of heat transfer. An analysis of the matter easily shows, [G4], that the correct[8] statement becomes exact only in the limit as $N, U \to \infty$, keeping of course $\frac{U}{V}, \frac{N}{V}$ constant, *i.e.* in what we call today the "thermodynamic limit".

Undoubtedly the word "holode" has the etymological origin of ὅλος and εἶδος while "ergode" is a shorthand for "ergomonode" and it has the etymological root of ἔργον and εἶδος, meaning a "monode with given energy", [G1]. [9]. The word "holode" is probably a shorthand for "holomonode",

[7] he also studies other ensembles, for instance in a system in which angular momntum is conserved, *e.g.* a gas in a spherical container, he considers the stationary distributions with fixed energy and fixed total angular momentum $\underline{L}$. Such monodes are called, by Boltzmann, *planodes* (form the "area law"); and he remarks that in general they are not orthodic (in fact one needs the extra condition that $\underline{L} = \underline{0}$).

[8] there is a problem only if one insists in defining in the same way the notion of heat transfer in the two cases: this is a problem that Boltzmann does not even mention, possibly because he saw as obvious that the two notion would become equivalent in the thermodynamic limit.

[9] the word "ergode" appears for the first time on p. 132 of [B84]: but this must be a curious misprint as the concept is really introduced on p. 134. On p. 132 the Author probably meant to say "holode", instead: this has been correctly remarked by [Pl]. See also footnote 13. The above etymology was probably proposed for the first time by myself in various lectures in Roma, and it was included in the first section of [G1]. The date of the preprint of [G1] is june 1980, the publication date is 1981: a year later a reference to the same new etymology appears, see [Ja],[Ma], attributed to Mathieu. I

meaning a "global monode" (perhaps a monode involving states with arbitrary energy, *i.e.* spread over the whole phase space).

This is not what is usually believed to be the etymology of "ergode": the usual belief comes from the Erhenfests' statement that the etymology is ἔργον and ὁδός, with the meaning of "unique path on the surface of constant energy, see [EE] note #93. This absurd etymology has been taken up universally and has been attached to the subject of "ergodic theory", which is instead a theory dealing with time evolution properties.

§2 *The ergodic hypothesis, continuous and discrete phase space.*

The etymological error of the Erhenfests could be just an amusing fact: but it had a rather deep negative influence in the development of the 20-th century Physics. They present their etymology in connection with the discussion (amounting to a *de facto* rejection) of the ergodic hypothesis of Boltzmann. In fact Boltzmann had come to the ergodic hyptothesis in his attempts to justify, *a priori*, that the ergode, as a model of thermodynamics, had to produce **the** thermodynamics of a system with the given hamiltonian function, (and not just a model).

Boltzmann had argued that the trajectory of any initial datum evolves on the surface of constant energy, visiting all phase space points and spending equal fractions of time in regions of equal Liouville measure.

The Ehrenfests criticize such a viewpoint on surprisingly abstract mathematical grounds: basically they say that one can attach to each different trajectory a different label, say a real number, thus constructing a function on phase space constant on trajectories. Such a function would of course have to have the same value on points on the same trajectory (*i.e.* it would be a constant of motion). This is stated in the note #74, p. 86 where the number of different paths is even "counted", and referred to in the note #94, p. 89. Therefore, they conclude, it is impossible that there is a single path on the surface of constant energy, *i.e.* the ergodic hypothesis is inconsistent (except for the monocyclic systems, for which it trivially holds).[10]

find it obviously possible, even likely, that independently two scientists may reach the same conclusion: even with only a few years of delay. Nevertheless *no reference* is made to my book in the paper of 1988 by Mathieu, in [Ma]. In fact I gave a series of lectures in august 1979 in Cortona which were attended by prof. R. Nagel who had access to (and, as all the partecipants, a copy of) my manuscript [G1] already including the etymology section in its present form; he informed me in a subsequent letter that he had discussed the matter with his student Mathieu, sending me a manuscript by him on the subject.

[10] the abstract mathematical nature of this argument, see also below for a critique, was

Having disposed of the ergodic hypothesis of Boltzmann, the Erhenfests proceed to formulate a new hypothesis, the rather obscure (and somewhat vague as no mention is made to the frequency of visit to regions in phase space) "quasi ergodic hypothesis", see notes #98 and #99, p.90, in [EE]: it led the physicists away from the subject and it inspired the mathematicians to find the appropriate definition giving birth to ergodic theory and to its first non trivial results.

The modern notion of ergodicity is not the quasi ergodicity of the Erhenfests. It is simply based on the remark that the Erhenfests had defined a non trivial constant of motion very abstractly, by using the axiom of choice. In fact from the definition, consisting in attaching a different number, or even $6N-2$ different numbers, to each distinct trajectory, there is *in principle* no way to construct a table of the values of the function defined in order to distinguish the different trajectories. In a system ergodic in the modern sense the Ehrenfests' construction would lead to a non measurable function; and to a physicist dowed with common sense *such a function, which in principle cannot be tabulated, should appear as non existent, or as non interesting.* Thus the motion on the energy surface is called ergodic if there are no *measurable* constants of motion: here measurable is a mathematical notion which essentially states the possibility of a tabulation of the function.

It is surprising that a generation of physicists could be influenced (in believing that the ergodic hypothesis of Boltzmann had to be abandoned as a too naive viewpoint) by an argument of such an exquisitely abstract nature, resting on the properties of a function that could not be tabulated (and not even defined if one did not accept the sinister axiom of choice).[11]

Therefore it is worth, perhaps, to try understanding what could have possibly meant Boltzmann when he formulated the ergodic hypothesis. Here one cannot fully rely on published work, as the question was never really directly addressed by Boltzmann in a critical fashion (he might have thought, rightly, that what he was saying was clear enough). The following analysis is an elaboration of [G1], [G2]: in some points it gets quite close to [Pl]. It will not escape the reader that [Pl] has a somewhat different point of

---

apparently remarked only by a mathematician, see [Pl] p. 86, althoug a great one (Borel, 1914); but it escaped many physicists. It is worrying to note how seriously the mathematicians took the ergodic hypothesis and how easily they disposed of it, taking for granted that the Ehrenfests formulation was the original formulation by Boltzmann and Maxwell, see [Br1],p. 383.

[11] we recall, as it is quite an irony, the coincidence that the recognition and the development of the axiom of choice was due essentially to the same Zermelo who was one of the strongest opponents of Boltzmann ideas on irreversibility, see also [Sc].

view on several key issues, although we seem to share the main thesis that the [EE] paper is responsible for most of the still persisting misunderstandings on Boltzmann's work. Including the exclusive attribution to Gibbs of Boltzmann's ideas on ensembles, so clearly elaborated in [B84].

My point of view is that of those who believe that Boltzmann always conceived the phase space as a discrete space, divided into small cells, see [B72], p. 346. He always stressed that the continuum must be understood as a limit, see [Br], p. 371, and [Kl1,2,3],[D]. The book of Dugas, [D], is particularly illuminating (also) on this respect (see for instance ch. 1 and the quotations of Boltzmann presented there, where he seems to identify the discrete viewpoint with the atomistic conceptions).

Although Boltzmann seems to have been, sometimes, quite apologetic about such a viewpoint (even calling it a "mathematical fiction", [Ba], p.18, from [B72]; see also [Pl], p. 75), he took advantage of it to a point that one can say that most of his arguments are based on a discrete conception of phase space, followed at the end by a passage to the continuum limit. It should be however understood that the discretization that Boltzmann had in mind is by no means to be identified with the later concept of coarse graining: see §4 where a modern version of Boltzmann's discretization is considered and where a distinction has to be made between cells and volume elements, see also [Pl] and [G3].

It is easier for us, by now used to numerical simulations, to grasp the meaning of a cell: in the numerical simulations a cell is nothing else but an element of the discrete set of points in phase space, each represented within computer precision (which is finite). One should always discuss how much the apparently harmless discreteness of the phase space affects the results. This is, however, almost never attempted: see [G3] for an attempt. A volume element has, instead, a size much larger than the machine resolution, so that it looks a continuum (for some purposes).

Hence one can say that an essential characteristics of Boltzmann's thought is to have regarded a system of N atoms, or molecules, as described by a *cell* of dimension δx and δp in each position and momentum coordinates. He always proceeded by regarding such quantities as very small, avoiding to enter into the analysis of their size, but every time this had some importance he must have regarded them as positive quantities.

A proof of this is when he refutes the Zermelo's paradoxes by counting the number of cells of the energy surface of $1cm^3$ of normal air, [B96], a feat that can only be achieved if one considers the phase space as discrete.

In particular this point of view must have been taken when he formulated the ergodic hypothesis: in fact conceiving the energy surface as discrete makes it possible to assume that the motion on it is "ergodic", *i.e.* it visits *all* the phase space points, compatible with the given energy (and possibly

with other "trivial" constants of motion) behaving as a monocyclic system (as all the motions are necessarily periodic).

The passage to the continuum limit, which seems to have never been made by Boltzmann, of such an assumption is of course extremely delicate, and it does not lead necessarily to the interpretation given by the Erhenfests. It can easily lead to other interpretations, among which the modern notion of ergodicity: but it should not be attempted here, as Boltzmann himself did not attempt it.

And in general one can hardly conceive that by studying the continuum problem could lead to really new information, that cannot be obtained by taking a discrete viewpoint. Of course some problems might still be easier if studied in the continuum, [S]: and the few results on ergodicity of physical systems do in fact rely explicitly on continuum models. However I interpret such results rather as illustrations of the complex nature of the discrete model: for instance the ergodicity theory of a system like a billiards is very enlightening as it allows us to get some ideas on the question of whether there exist other ergodic distributions (in the sense of ergodic theory) on the energy surface, and which is their meaning, [BSC].

And the theory of the continuum models has been essential in providing new insights in the description of non equilibrium phenomena, [R], [CELS].

Finally the fruitfulness of the discrete models can be even more appreciated if one notes that they have been the origin of the quantum theory of radiation: it can be even maintained that already Boltzmann had obtained the Bose Einstein statistics, [Ba].

The latter is a somewhat strong intepretation of the 1877 paper, [B77]. The most attentive readers of Boltzmann have, in fact, noted that in his discretizations he really thinks always in terms of the continuum limit as he does not discuss the two main "errors" that one commits in regarding a continuum formulation as an approximation (based on integrals instead of sums)[12] with respect to a discrete one.

The above "oversight" might simply be a proof that Boltzmann never took the discretization viewpoint to its extreme consequences. Among which there is that the equilibrium ensembles are *no longer* orthodic in the sense of Boltzmann (see [G3],[G4]), (although they still provide a model for thermodynamics provided the temperature is no longer identified with the average kinetic energy): a remark that very likely was not made by Boltzmann *in spite of his consideration and interest on the possibility of finding other integrating factors for the heat transfer* dQ, see the footnote

[12] and which amount to the identification of the Maxwell Boltzmann statistics and the Bose Einstein statistics, and to neglecting the variation of physically relevant quantities over the cells: see the lucid analysis in [K], p.60; for a technical discussion see [G3],[G4].

on p. 152 in [B84].[13]

The necessity of an understanding of this "oversight" has been in particular clearly advocated by Kuhn referring to Boltzmann's "little studied views about the relation between the continuum and the discrete", [K], for instance.

§3 *The ergodic hypothesis and irreversibility.*

The reaction of the scientific world to the ergodic hypothesis was, "on the average", a violently negative one, also as it was intended to provide further justification to the irreversibility predicted by the Boltzmann equation, derived earlier.

The great majority of the scientists saw absurd and paradoxical consequences of the hypothesis, without apparently giving any importance to the "unbelievable" fact that on the basis of a maximal simplicity assumption (*i.e.* only one cycle on the energy surface) Boltzmann was obtaining not only the possibility of explaining, mechanically, the classical equilibrium thermodynamics but also that of explaining it in a quantitative way. It allowed, for the first time, the theoretical calculation of the equations of state of many substances (at least in principle) like imperfect gases, and even other fluids and solids.

The success of the highly symbolic but very suggestive formula of Boltzmann, see [EE], p.25:

$$\lim_{T\to\infty} \frac{dt}{T} = \frac{\sigma\, ds}{\int \sigma\, ds} \tag{3.1}$$

(where σ is the microcanonical density on the energy surface, whose area element is ds) in the calculation of the equilibrium properties of matter led quickly the physicists to accept it in the "minimal interpretation". Such interpretation demanded that the r.h.s. be used to compute the equilibrium averages and the l.h.s. ignored, togheter with the atomic hypothesis. This is regarded as a *law of nature*, in spite of the persistent skepticism (or deep doubts) on its deducibility from the laws of mechanics. A point of view usually attributed to Gibbs, referring to [G], and which is still around us, although we assit, since the mid fifties, to a slow but inexorable inversion of tendency.

Immediately after the first critiques Boltzmann elaborated answers often very clear and simple by our modern understanding: but they were very

[13] I have profited, in checking my understanding of the original paper as partially exposed in [G1], from an english translation that Dr. J. Renn kindly provided, while being ny student in Roma (1984). I could note this footnote in [B94], and insert a few new remarks in the present paper, because of his translation, (unfortunately still unpublished).

frequently ill understood not only by the opponents of Boltzmann and their epigones, but also by those who were closest to him. The above quoted critique to the ergodic hypothesis by the Erhenfests is a shocking example.

Another example is the recurrence paradox, based on the simple theorem of Poincaré. Boltzmann was finally led to the calculation of the number of cells on the energy surface, [B96], thus to a superastronomical estimate of the recurrence time: which, nevertheless, did not seem to impress many.

It is also clear that Boltzmann himself became aware of the fact that, after all, the ergodic hypothesis might have been unnecessarily strong and perhaps even useless to explain the approach to equilibrium in physical systems. The latter in fact reach equilibrium, normally, within times which are microscopic times, not at all comparable with the recurrence time. He asserted repeatedly that the (very few) macroscopic observables of interest had essentially the *same* value in most of the energy surface, and the time spent in the "anomalous phase space cells" is therefore extremely small: a quantitative understanding of this is provided by the Boltzmann equation. This remark also frees (3.1) from the ergodic hypothesis: it might well be that the r.h.s can be used to evaluate the average values, in equilibrium, of the few observables which are of interest, although there might be observables (*i.e.* functions on phase space) for which the (3.1) fails.

It is well known that Boltzmann went quite far in this direction, by providing us with a concrete method to estimate the true times of approach to equilibrium: the Boltzmann's equation (historically developed well before the 80's).

Finally it is worth noting that the methods used by By Boltzmann in deriving the theory of the ensembles and the ergodic hypothesis are quite modern and in fact are most suited to illustrate the new developments on non equilibrium theory: as I shall try to prove in the next section.

§4 *Non equilibrium. Ruelle's principle. Outlook.*

I cannot resist the temptation of at least mentioning some recent new developments which look to me exciting and very likely to remain as important progress in the field.[14]

The (3.1), in its minimal interpretation of providing, via the r.h.s. (*i.e.* the microcanonical distribution), the law for the evaluation of the "relevant"

[14] I like to think that Boltzmann his listening to the celebration of his birthday: he would certainly be bored by hearing a, presumably poor, exposition dealing only with things that he knew far better.

macroscopic observables, starting from the energy function of the system, "solves" the problem of the equilibrium theory. Completely, as far as we know (in Classical Physics).

Is a similar theory possible for systems in non equilibrium, but in a stationary state? What (if anything) replaces the microcanonical distribution in such cases? As an example of "cases" we mean the motion of a gas of particles subject to a constant force ("electric field") setting them in motion, while the energy produced is dissipated into a reservoir.

The answer seems positive, at least in some cases. The problem lies in the fact that the motion of such systems is dissipative, hence the volume element of the energy surface is not conserved even in the simple case in which the thermostat is such that it keeps the total energy of the system constant (as I shall suppose, to simplify the discussion), *i.e.* the microcanonical distribution cannot describe the stationary state. Taking the continuum viewpoint we can imagine that the motion is essentially concentrated, after a transient time, on a set A which has zero measure with respect to the Liouville measure on the energy surface.

To avoid giving the impression that the discussion is abstract (hence possibly empty) let me declare explicitly one, among many, models that one should have in mind. We consider a system of N particles interacting with a potential energy Φ and subject to an external constant force field $\underline{E}$, (*e.g.* electric field):

$$\dot{\underline{q}}_i = \frac{1}{m}\underline{p}_i, \qquad \dot{\underline{p}}_i = -\partial_{\underline{q}_i}\Phi + \underline{E} - \alpha(\underline{p})\underline{p}_i \tag{4.1}$$

where $\underline{E}$ is the external constant force and α is defined so that the energy $\sum_{i=1}^{N}\frac{\underline{p}_i^2}{2m} + \Phi$ is constant (*i.e.* $\alpha = \frac{\underline{E}\cdot\sum\underline{p}_i}{\sum\underline{p}_i^2}$). The term $\alpha\underline{p}_i$ is a model of a thermostat (this should be called a *gaussian thermostat* as it is related to the Gauss' principle of "least constraint", see [CELS]). The system is considered enclosed in a box with periodic boundary conditions: hence we expect that a current parallel to $\underline{E}$ will be established and the system will reach a stationary state. The volume in phase space contracts at a rate $(3N-1)\alpha$, (which is positive, in the average): hence the motion will asymptotically develop on some "attractor", which is a set of 0 Liouville measure.

What follows will lead to a unified theory of the equilibrium as well as the non equilibrium, for system (4.1).

The discrete viewpoint is also possible: the enegy surface consists of cells which are relevant (for the study of the asymptotic properties) forming a set A in phase space, and of cells which are irrelevant. The motion can be regarded to develop on the set of cells which are in A, which is strictly

smaller than the set of all the cells: in fact far smaller (and in the continuum limit the fraction of cells in A approaches 0).

Since the volume of the cells is not conserved care must be exercised in regarding the dynamics as a permutation of the cells of A. This is in fact also true in the equilibrium case because, even if the cells do not change in volume, they are deformed being squeezed in some directions and dilated in others. In equilibrium it is possible to conceive situations in which the deformation can be neglected (this leads to restrictions on the region of temperature and density in which the consideration of the dynamics as a cell permutation is acceptable: a discussion which we have not begun above and which we avoid here as well, see [G3] for a quantitative analysis). And a similar analysis can be carried in the present case.

Basically one has to think that the system is observed at time intervals τ_0 which are not too small (so that something really happens) and not too large (so that the cell's deformations can be either neglected or controlled, at least for a large majority of cells): see [G3] for a quantitative analysis of what this means in the equilibrium cases and of when this might lead to inconsistencies. Let S_{τ_0} denote the transformation of A describing the dynamics on A over the time τ_0. By making the cells small enough we can take τ_0 larger.

We shall imagine the set A as a surface in phase space of dimension roughly $\frac{6N}{2}$ at least if the external force is small (so that the friction α, *i.e.* the phase space volume contraction, is also small): in fact if there is no external force the dimension of A should be $1 + \frac{6N-2}{2}$.[15] The surface A can fold itself on the energy surface filling it up completely (in the $\underline{E} = \underline{0}$ case) or not (in the general case).[16] We can assume the following extension of the ergodic hypothesis: *on A the dynamics is a one cycle permutation of the cells.*

[15] because there are as many contracting directions as expanding ones (the volume being conserved in the $6N$ dimensional phase space); and there are two "neutral" directions (the direction orthogonal to the energy surface and the direction of the phase space motion) one of which lies on the energy surface (the direction of motion), see [Dr], [ECM1], [SEM]. Of course the existence of other conserved quantities, as in (4.1) when the linear momentum is conserved, affects this calculation: in (4.1), when $\underline{E} = \underline{0}$, this brings down the dimension to $1+\frac{6N-8}{2}$. Furthermore we are assuming here that there are no "neutral" directions other than the ones possibly provided by the obvious conservation laws: *i.e.* that our system has strong instability properties (hence this does not *directly* apply to the free gas, for instance).

[16] in the continuum point of view we can proceed as follows: we fix an approximation ε and we identify the points on A which are very far on any path that joins them *along* A, but which are close within ε as points on the energy surface. Then A becomes a finite surface A_ε. This surface depends on the point that we initially choose for the construction: but the results should be independent on the choice. The latter is in fact an assumption

Then the motion of a randomly chosen initial datum, randomly with respect to a distribution with some density on the energy surface, will simply consist in a fast approach to the surface A; at the same time data which are on A itself and close to each other will separate from each other at some exponential rate, because on A all the directions are dilated, by definition. To fix the ideas we take the initial data with constant density in some little ball U. If we assume, for simplicity, the above ergodic hypothesis, the layer is, over times multiples of the recurrence time, a set of cells each visited with equal frequency. However the surface A will, in general, not be a monolayer of cells but it will have a large "width", *i.e.* a (macrospcopic) area element $d\sigma$ will contain many (microscopic) cells.[17]

The number of cells per unit area can be deduced by remarking that after a time $\tau = M\tau_0$ the density of cells around $x \in A$, initially distributed with constant density in the region U (where the initial data are randomly chosen), has to be proportional to the inverse of the area expansion rate of the transformation S_τ. This means that we expect that the distribution on A which has to be used to compute the stationary averages is described by a suitable density with respect to the area element on A.

With this intuitive picture in mind, [R], ECM2], we see that a little ball U in phase space evolves becoming a thin layer around A: the density of the layer, after a large time T, is proportional to the expansion rate of the surface area on A under the transformation S_T generating the time evolution over the given time.

In the case of no external forces one has that the surface A folds itself on the energy surface coming back to a given phase space volume element V_0 (not to be confused with a cell, which has to be thought as much smaller); just enough times, and with enough volume around, so that the fraction of the volume initially in U and falling in the volume element V_0 is proportional to V_0 itself (this is consistent because of the equality of the total expansion rate and the total contraction rate, due to the hamiltonian na-

---

which essentially replaces the ergodicity assumption of the conservative cases. The above "viewpoint" will imply ergodicity in the case of the conservative systems: this non trivial fact is a consequence of the hidden assumption that the description does not depend on which surface A_ε we choose as an approximation for A. In fact the choice of A_ε suffers from an arbitrariness which consists in deciding that one given point is actually on A_ε: choosing another point leads, in general, to a different A_ε. In concrete cases it will, however, be very difficult to show that the results are independent on A_ε (a manifestation of the conservation of difficulties).

[17] this can perhaps be clarified if one thinks of the numerical experiments in which the computer representatives of the phase space points are regarded as cells, while the unstable manifolds of the motion are regarded as surfaces built with computer points, *i.e.* cells.

ture of the equations of motion). But in general the fraction of volume U falling into a volume element will be far different from the volume element fraction of the energy surface.

One is thus led to the following unified "principle" to describe the stationary states of non equilibrium systems, [R]:

Principle: *the average values of the observables in the stationary state describing the asymptotic behaviour of systems like (4.1), is computable from a probability distribution on A which has a density, with respect to the surface element of A.*[18]

This principle can be more mathematically stated (a problem into which we refrain to enter here), and is due to Ruelle, [R], who based himself also on the results of Anosov, Sinai, Bowen on the theory of a class of dynamical systems known as "hyperbolic systems" (which play in some sense, for non equilibrium statistical mechanics, the role of the monocyclic systems of Helmholtz). The probability distributions selected by the above principle (which in "good cases" is unique) are called SRB measure, [R].

What is the predictive value of the above statements? in the cases without external forces we have already mentioned that this principle leads to the microcanonical distribution and, therefore, implies the classical thermodynamics, [B84]. Life is made easy by the fact that although A may be very difficult to identify, still the stationary distribution is just the microcanonical ensemble because A folds on the energy surface filling it up completely, with no gaps.

In the dissipative cases it seems that we have little control on A and hence on the stationary distribution.

Yet this might not be really so: we simply have to learn how to extract informations from such an abstract principle. After all it now seems natural that the Gibbs distribution predicts all the phenomena of equilibrium statistical mechanics (from the phase coexistence, to the critical point, to cristallization). But this was far from clear only a few decades ago, and many decades after the original formulations of Maxwell, Gibbs and Boltzmann (as many of us certainly recall).

That the principle might have predictive value is indicated by the first attempts at its use in problems of statistical mechanics, see [ECM2], (see also [CELS]), who were somewhat inspired by previous papers, see also [HHP]. In fact only recently the principle started being considered in the theory

[18] it is extremely important to think, to avoid trivial contradictions, that the cells on A must be regarded as much smaller than the surface elements of A that we consider in talking about the density.

of non equilibrium, as it was developed originally by Ruelle mainly as an attempt to a theory of turbulent phenomena. This is not the appropriate place to discuss the !xpaper [ECM2] in the perspective of the above principle: the discussion is rather delicate (as [ECM2] should be regarded as a pioneering work).

A simpler example of a quantitative (yet quite abstract) consequence of the above principle is the determination of the density function mentioned in the principle: the latter is in fact essentially determined. If we are interested in stationary distributions phenomena which are observable by measurements that take place in a fixed time τ we can just take averages over A with respect to a distribution with density over A proportional to $\Lambda_{\tau'}^{-1}(x)$, with $\tau' = M'\tau_0 \gg \tau$ (where the expansion rate is the jacobian determinant of the transformation $S_{\tau'}$ at x, *i.e.* $\Lambda_{\tau'}^{-1}(x) \equiv \prod_{-M'}^{M'} \Lambda_{\tau_0}^{-1}(S_{\tau_0}^j x)$). So that two equal area elements of A around x and y have a *relative probability* of visit equal to $\Lambda_{\tau'}^{-1}(x)/\Lambda_{\tau'}^{-1}(y)$.

Of course τ' cannot be taken too large: if τ' is taken of the order of the recurrence time the ratio becomes 1. The natural upper bound on τ' has to be such that the cells in U ending in the considered area elements are still in a large number. This sets an upper limit to the values of τ for which the above remark applies. [19]

The example (4.1) is very special.[20]

It is however generalizable: many generalizations have already been considered in the literature, [PH]. Still it should be stressed that the models to which the above principle can be applied form a rather small class of deterministic models. It is not immediately clear how it can be applied to stationary non equilibrium phenomena in which the thermostat is realized in a different way, *e.g.* by some stochastic boundary conditions. Nor it is obvious that the different thermostats are physically equivalent.

In my opinion there is, also, some misunderstanding in the literature about the fact that the set A has zero measure (in the non equilibrium cases this has been sometimes associated with the questions related to irreversibility) and about the fact that A, regarded as a folded surface on the equal energy

[19] this means that the ratio between the linear dimension of U and the linear dimension of the cells has to be large compared to the maximal linear expansion rate over the time τ, a condition that can be expressed in terms of the largest Lyapunov exponent.

[20] this is shown also by the fact that the operation i mapping $x = (\underline{p}, \underline{q})$ to $ix = (-\underline{p}, \underline{q})$ is such that $t \to ix(-t)$ is a solution of the equation of motion if $t \to x(t)$ is such: a time reversal symmetry. This has several implications, among which the properties that both initial data x and ix evolve towards the same attractor A, in the future, and to the attractor iA in the past. In general A and iA are different, except in the case $\underline{E} = \underline{0}$ (because A is the full energy surface).

manifold, has a fractal dimension (thereby representing a "strange attractor"). Such facts may be quite misleading. The above analysis shows that A should be more conveniently regarded as a smooth non fractal surface of dimension about $6N/2$: its fractal dimension arises from the folding of A on the surface of constant energy (rising from $6N/2$ to about $6N$ if $\underline{E}$ is small).

Furthermore in the assumption that the stochastic thermostats and the gaussian thermostat (or other thermostats, [PH]) are equivalent one sees clearly a problem related to attaching importance to the set A as a fractal with zero measure. In fact we expect that stochastic thermostats lead to stationary distributions which have a density in phase space, hence which cannot be concentrated on a set of 0 measure.

The contradiction disappears if one thinks that, in a stationary state, there may be several distributions which, in the limit as $N \to \infty$, become equivalent. A distribution concentrated on a set of zero measure might well be equivalent to one distributed on the whole energy surface, or on the whole phase space. A much simpler, but very familiar, example of such a situation is provided by the microcanonical distribution which is concentrated on a set of zero measure, but it is equivalent (in the thermodynamic limit) to the canonical distribution, which is concentrated on the whole phase space.

Finally it should be clear that the problem of approach to stationarity will show up exactly in the same terms as in the equilibrium cases. The "ergodicity" assumptions above cannot in any way justify the use of the distribution verifying the Ruelle principle: the time necessary for a phase space point to visit the full set of cells building A will be of the order of magnitude of the recurrence time. And as in the equilibrium cases we can expect that the rapidity of the approach to equilibrium is rather due to the fact that we are interested only in very few observables, and such observables have the same value in most of phase space.

I hope to have shown, or at least given arguments, that the point of view, see for instance [Pl], whereby Boltzmann was a XIX century physicist judged by his interpreters with XX century mathematical standards is not exactly correct: today's way of thinking is not too different from his and most problems the physicists had with his work were problems with the understanding of his Physics and *not* of his Mathematics, see also [L]. The misunderstandings about his ideas are, in my opinion, largely due to the unwillingness of studying the original publications and to the unfounded belief that they were forwarded with fidelity by the reviewers that wrote about his achievements.

Acknowledgements: I owe to my father Carlo essential help in the ex-

planation of the etymology of the word ergodic. Part of the interpretation of Ruelle's principle presented here was developed in collaboration with E. Cohen in a joint effort to understand more deeply the results of the paper [ECM2]: while our analysis, which preceded this paper, will be published elsewhere I wish to thank him for communicating to me his enthousiasm on the subject while I was visiting Rockefeller University, and for his thoughtful comments on this paper. I am indebted to J. Lebowitz for his hospitality at Rutgers university and for stimulating my interest on the gaussian thermostats. To him I owe also the redressement of several misconceptions and mathematical errors.

References.

[B66] Boltzmann, L.: *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie*, in "Wissehschaftliche Abhandlungen", ed. F. Hasenhörl, vol. I, p. 9–33, reprinted by Chelsea, New York).

[B68] Boltzmann, L.: *Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten*, in "Wissehschaftliche Abhandlungen", ed. F. Hasenhörl, vol. I, p. 49–96, reprinted by Chelsea, New York.

[B72] Boltzmann, L.: *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*, english translation in S. Brush, *Kinetic theory*, Vol. 2, p. 88. Original in "Wissehschaftliche Abhandlungen", ed. F. Hasenhörl, vol. I, p. 316–402, reprinted by Chelsea, New York).

[B77] Boltzmann, L.: *Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respektive den Sätzen über das Wärmegleichgewicht*, in "Wissenschaftliche Abhandlungen", vol. II, p. 164–223, F. Hasenöhrl, Chelsea, New York, 1968 (reprint).

[B84] Boltzmann, L.: *Über die eigenshaften monzyklischer und anderer damit verwandter Systeme*, in "Wissenshafltliche Abhandlungen", ed. F.P. Hasenhörl, vol. III, Chelsea, New York, 1968, (reprint).

[B96] Boltzmann, L.: *Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo*, english translation in S. Brush, "Kinetic Theory", vol. 2, 218–, Pergamon Press.

[B97] Boltzmann, L.: *Zu Hrn. Zermelo's Abhandlung "Ueber die mechanische Erklärung irreversibler Vorgänge*, english translation in S. Brush, "Kinetic Theory", **2**, 238.

[B02] Boltzmann, L.: *Lectures on gas theory*, english edition annotated by S. Brush, University of California Press, Berkeley, 1964.

[Ba] Bach, A.: *Boltzmann's probability distribution of 1877*, Archive for the History of exact sciences, **41**, 1-40, 1990.

[Br1] Brush, S.: *The kind of motion we call heat*, North Holland, 1976 (vol. II), 1986 (vol. I).

[BSC] Bunimovitch, L., Sinai, Y., Chernov, N: *Statistical properties of two dimensional hyperbolic billiards*, Russian Mathematical Surveys, **45**, n. 3, 105–152, 1990.

[C] Clausius, R.: *The nature of the motion which we call heat*, in "Kinetic Theory, ed. S. Brush, p. 111—147.

[CELS] Chernov, K., Eyink, G., Lebowitz, J., Sinai, Y.: *Steady state electric conductivity in the periodic Lorentz gas*, Communications in Mathematical Physics, **154**, 569–601, 1993.

[D] Dugas, R.: *La théorie phisique au sens de Boltzmann*, Griffon, Neuchâtel, 1959.

[Dr] Dressler, U.: *Symmetry property of the Lyapunov exponents of a class of dissipative dynamical systems with viscous damping*, Physical Review, **38A**, 2103–2109, 1988.

[ECM1] Evans, D.,Cohen, E., Morriss, G.: *Viscosity of a simple fluid from its maximal Lyapunov exponents*, Physical Review, **42A**, 5990–5997, 1990.

[ECM2] Evans, D.,Cohen, E., Morriss, G.: *Probability of second law violations in shearing steady flows*, Physical Review Letters, **71**, 2401–2404, 1993.

[EE] Ehrenfest, P., Ehrenfest, T.: *The conceptual foundations of the statistical approach in Mechanics*, Dover, 1990, (reprint).

[G] Gibbs, J.: *Elementary principles in statistical mechanics*, Ox Bow Press, 1981, (reprint).

[G1] Gallavotti, G.: *Aspetti della teoria ergodica qualitativa e statistica del moto*, Quaderni dell' U.M.I., vol. 21, ed. Pitagora, Bologna, 1982.

[G2] Gallavotti, G.: *L' hypothèse ergodique et Boltzmann*, in "Dictionnaire Phylosophique", Presses Universitaires de France, p. 1081– 1086, Paris, 1989.

[G3] Gallavotti, G.: *Meccanica Statistica*, entry for the "Enciclopedia italiana delle scienze fisiche", preprint Roma, 1984. In print (scheduled publication, 1994). The published version will also include another entry, originally written to be a separate one, *Equipartizione e critica della Meccanica Statistica Classica*, Roma, preprint 1984. See also the entry *Teoria Ergodica*, preprint Roma, 1986, for the "Enciclopedia del Novecento", (in print? maybe).

[G4] Gallavotti, G.: *Insiemi statistici*, entry for the "Enciclopedia italiana delle scienze fisiche", preprint Roma, 1984. In print (scheduled publication, 1994).

[He1] Helmholtz, H.: *Principien der Statik monocyklischer Systeme*, in "Wissenschaftliche Abhandlungen", vol. III, p. 142–162 and p. 179– 202, Leipzig, 1895.

[He2] Helmholtz, H.: *Studien zur Statik monocyklischer Systeme*, in "Wissenschaftliche Abhandlungen", vol. III, p. 163–172 and p. 173– 178,

Leipzig, 1895.

[HHP] Holian, B., Hoover, W., Posch. H.: *Resolution of Loschmidts paradox: the origin of irreversible behaviour in reversible atomistic dynamics*, Physical Review Letters, **59**, 10–13, 1987.

[Ja] Jacobs, K.: *Ergodic theory and combinatorics*, in Proceedings of the conference on *Modern analysis ansd probability*, june 1982. Contemporary Mathematics, **26**, 171–187, 1984.

[K] Kuhn, T.: *Black body theory and the quantum discontinuity. 1814–1912*, University of Chicago Press, 1987.

[Kl1] Klein, M.: *Maxwell and the beginning of the Quantum Theory*, Archive for the history of exact sciences, **1**, 459–479, 1962.

[Kl2] Klein, M.: *Mechanical explanations at the end of the nineteenth century*, Centaurus, **17**, 58–82, 1972.

[Kl3] Klein, M.: *The development of Boltzmann statistical ideas*, in "The Boltzmann equation", ed. E. Cohen, W. Thirring, Acta Physica Austriaca, suppl. X, Wien, p. 53–106.

[L] Lebowitz, J.: *Boltzmann's entropy and time's arrow*, Physics Today, Sept 1993, p. 32–38.

[LPR] Livi, R., Politi, A., Ruffo, S.: *Distribution of characteristic exponents in the thermodynamic limit*, Journal of Physics, **19A**, 2033–2040, 1986.

[M] Maxwell, J.: *On Boltzmann's theorem on the average distribution of energy in a system of material points*, in "The scientific papers of J,C, Maxwell", ed. W. Niven, Cambridge University Press, 1890, vol. II, p. 713–741.

[Ma] Mathieu, M.: *On the origin of the notion 'Ergodic Theory'*, Expositiones Mathematicae, **6**, 373–377, 1988. *See footnote* [9] *above.*

[Pl] Plato, J.: *Boltzmann's ergodic hypothesis*, Archive for the History of exact sciences, **44**, 71-89, 1992.

[H] Posch, H., Hoover, W.: *Non equilibrium molecular dynamics of a classical fluid*, in "Molecular Liquids: new perspectives in Physics and chemistry", ed. J. Teixeira-Dias, Kluwer Academic Publishers, p. 527–547, 1992.

[R] Ruelle, D.: *Measures describing a turbulent flow*, Annals of the New York Academy of Sciences, **357**, 1–9, 1980. See also Eckmann, J., Ruelle, D.: *Ergodic theory of strange attractors*, Reviews of Modern Physics, **57**, 617–656, 1985; and Ruelle, D.: *Ergodic theory of differentiable dynamical systems*, Publications Mathémathiques de l' IHES, **50**, 275–306, 1980.

[S] Sinai, Y.: *Dynamical systems with elastic reflections. Ergodic properties of dispersing billards*, Russian Mathematical Surveys, **25**, 137–189, 1970.

[Sc] Schwartz, J.: *The Pernicious Influence of Mathematics on Science*, in "Discrete thoughts: essays in Mathematics, Science, and Phylosophy", M. Kac, G. Rota, and J. Schwartz, eds., Birkhauser, Boston, 1986, p. 19–25.

[SEM] Sarman, S., Evans, D., Morriss, G.: *Conjugate pairing rule and thermal*

transport coefficients, Physical Review, **45A**, 2233–2242, 1992.

Qualche lettore potrebbe trovare interessante il seguente rapporto del referee (da me ricopiato senza modifiche) e forse anche la mia risposta (che lo segue). Le revisioni apportate in conseguenza della relazione del referee, e di altri commenti ricevuti, hanno carattere poco rilevante, a mio parere; assai piú interessante e utile può essere la lettura dei commenti assai negativi del referee. Pertanto, per correttezza verso di lui, la versione sopra riportata del lavoro non e' la versione pubblicata sul Journal of Statistical Physics, bensì la versione che originariamente fu inviata al comitato di redazione di quella rivista. Questa versione è archiviata su *mp_arc@math.utexas.edu* con il numero 93-66 e su $chao-dyn@xyz.lanl.gov$ con il numero 9403004. Una correzione rilevante che ho apportato, grazie ad un commento ricevuto, è la variazione dell' estremo superiore del prodotto nelle prime righe di p.214 da M' al valore corretto 0.

COMMENT FROM THE REFEREE (anonymous):

As you will see, my review though lengthy and negative on the first parts of the paper, does not deal with Section 4, where Gallavotti presents recent work in the light of Boltzmann's ideas. I am not competent to criticize this, but in my suggestions to the author I urge him to put the emphasis there.

Referee's Report JSP 94-73, ms by G. Gallavotti*

I must begin by saying that Dr. Gallavotti does not make clear, at least not to this reader, just what he is trying to convey in this paper. This is due, at least in part, to his heavy stress on the etymology of the word "ergodic," and to the discussion he gives of the "etymological error of the Ehrenfests" and its "rather deep negative influence in the development of twentieth century physics." I shall make a suggestion about a possible reorganization of the paper at the end of my review.

To begin then, as Gallavotti does, with the etymology: I have never studied Greek so I cannot judge the issue directly. Note the following points, however. 1) Paul Ehrenfest was Boltzmann's student and moved in his circle for four years or so. As Boltzmann's prize student it seems likely that

* [the footnotes numbers refer to the original version and some are shifted by one or two in the present revised version to which a few footnotes have been added]

he had discussed ergodicity with his teacher. Ehrenfest, too, had studied at the Gymnasium, and had been taught Greek. 2) There certainly are terms in mechanics which are based on odos (path). Think of the hodograph, the polhode, and the herpolhode. Both Boltzmann and Ehrenfest knew mechanics very well. 3) It is probably quite irrelevant, but in this paper the name Ehrenfest is mis-spelled quite consistently as Erhenfest. Since the proper spelling has a perfectly good etymology while the one used here does not have even an "absurd" one, this reader is inclined to be skeptical of the whole business. In any case the etymology given has to be stretched a good deal to cover both holode and ergode (p. 6). The frequent use of the word "undoubtedly" also does not encourage confidence in the argument. (See note 5 on p. 3.)

Now to return to the real issues. I agree that the question posed in Boltzmann's 1884 paper is an important one, namely, for which stationary ensembles do the laws of thermodynamics apply to the ensemble averages? (This is the content of (1.1) and (1.2) with the additional condition that the average kinetic energy serve as the absolute temperature. This question was answered for the canonical ensemble (the holode) by Boltzmann in his 1871 paper "Analytischer Beweis...." Equilibrium ensembles as such are attributed to Gibbs only by those unaware of Boltzmann's papers. This certainly does not include Stephen Brush, and if p. 242 of his book is read carefully one sees that he correctly gives Gibbs proper credit for the grand canonical ensemble and only for that. (He does however seem to overlook Boltzmann's 1871 paper there.)

In section 2 of this paper Gallavotti claims that Boltzmann "always" conceived of a phase space made up of discrete cells. But then, he agrees, Boltzmann goes to the continuum as a limit. One could say the same about many of Newton's arguments or indeed about much of the calculus. After claiming that Boltzmann takes the discrete argument and representation as basic, on pp. 9-10, he remarks on p. 12 that "in his discretizations [Boltzmann] really thinks always in terms of the continuum limit," which sounds to me like the opposite of the earlier claim. Once again there are several "must have" phrases when Gallavotti is telling us about Boltzmann's thinking.

Section 3 begins with two historical assertions about how "the scientific world" or "the great majority of scientists" responded to the ergodic hypothesis. No evidence is presented for these assertions. Further, why is it appropriate in 1994 to think that the time average for a single system is what "really" counts, and the ensemble average is only an expedient for calculating it?

The last section of the paper deals with matters beyond my competence. Whether or not that unintelligibility is all my fault, I cannot say. (I do,

however, find it amusing that the author wants "to avoid giving the impression that [his] discussion is abstract," only a few lines before introducing a "gaussian thermostat," a quantity which has the dimension of inverse time and which is proportional to the constant force acting on the system.)

My advice to the author for the revisions that I consider necessary consists of the following points.

1.) Begin the paper by pointing out the positive content of Boltzmann's 1884 paper, using modern terminology. Put the etymology and Boltzmann's various "odes" into an -appendix-. But if you want to insist on your etymology and rule out the Ehrenfests', you need to have solid arguments, so far missing. (I would urge that the priority dispute in footnote 9 be omitted, handled privately, or at least be reduced to a sentence or two.)

2.) I suggest much less emphasis on the negative consequences of the Ehrenfests' shift in the meaning of ergodic. This point is discussed in some detail by Brush on pp. 363-377 of his book, "The Kind of Motion." These pages are in the second volume. I don't find much discussion of the ergodic problem by physicists at that time. The role of discreteness in Boltzmann's work is indeed interesting and important, but this discussion is partly conjectural and somewhat confusing.

3.) The last section of the paper contains the new and recent material. I think it should be made more central to the presentation with more discussion and something closer to real examples, if possible.

RISPOSTA:

Reply to the referee of the paper:
"Ergodicity, ensembles, irreversibility in Boltzmann and beyond."
Author: G. Gallavotti, submitted to J. Statistical Physics.

1) the paper is NOT based on my knowledge of greek. In fact, unfortunately, I do not know greek. Although I studied in the Gymnasium I was never bothered by the etymology of the word ergode until I decided to look at the original. Even then my rudiments of greek were not sufficient to realize that the etymology was incorrect: I was however bothered by the difficulty to match the meaning that was conveyed by the paper with the official etymology and the correct etymology was pointed out to me by my father (see acknowledgements) who explained me that there ought to have been an "h" in front of "ode" had it originated from "odos". I thought that my paper could be understood essentially without any reference to greek language.

2) Since the referee brings the point up I am forced, nevertheless, to remark that he correctly derives from "odos" the words he mentions. The proof is that they all have an h before the ending "ode". This is because in greek "odos" has an aspiration accent. Therefore I regard this comment as a strengthening of my etymology. But I did not want to give this argument (that I learnt from my father) as it is based on phylology and I myself did not notice anything linguistically wrong with the missing h: my argument is based on the actual contents of the paper of Boltzmann B84. To insist on phylological arguments would not be right: after all why should Boltzmann adhere to Erasmus transliteration rules? why should he know them and not be free to form the words he liked most, attributing them the meaning he wanted? But the paper B84 shows that things are different: I tried to discuss this and I cannot agree on the "stretching" that the referee sees in my analysis. I would, instead, very much like any criticism based on the Boltzmann paper contents: to receive such criticisms was one of the reasons for writing the paper.

3) I do not grasp whether the comment on the spelling of Ehrenfest is a joke or demands an answer. I take it as a polite way of pointing out a (particularly irrelevant but certainly annoying) error on my side and I correct it.

4) about the word "undoubtedly" I thought that the note 5 on p. 3 just says what the referee would like me to say, he seems to have noticed it but somehow he does not give to it the meaning I wanted it to have. Namely that the word "undoubtedly" is just a word which is used to express my strong opinion but at the same time it leaves room for alternative opinions (as the footnote states).

5) I thank the referee for the comment on Brush, p.242: I had misinterpreted that page. I have changed the reference to Brush referring to it in a different way here as well as in the other place of the paper where I was talking about it. However I cannot change completely my statement.

6) The referee's comment about the continuoum limit is not correct and I think he has been misled by my writing, admittedly unclear. The phrase in my paper is, with respect to the paper, very important and refers to a technical question. Namely by no means I agree that Boltzmann regards the space as a continuoum. I wanted to say that he (as well as anybody else) prefers to compute elementary integrals rather than the sums they stand for, either as a simpler symbol or as a DEVICE. After all the calculus courses, that are imposed on us, one is used to have recourse to well known rules (eg $\int x = x^2/2$, or $\int sinx = -cosx$, etc.). This happens only for a few trivial integrals: all the others are better regarded as sums (in France it is still very common to read the sign "integral" by calling it "sum"). He simply had no reason to put a bound on the cells size as for him that

would have had no consequences at the places where he was applying the procedure. The consequences of discreteness and its relevance in extreme situations was realized by Planck, much later: this point is also brought up in my paper in connection with the phrase that disturbs the referee. I have modified the phrase and I hope it is now clearer.

7) my "must have" bother the referee: they simply mean that I prefer to say that B. "must have thought the phase space as discrete" when he counted the number of cells in it, rather than to say that I did not understand why he did not simply say, right away, that the number was infinite and then he looked for a better answer to Zermelo recurrence paradox (discarding the one he gave, which I find beautiful). In fact I could take away the word "must", as I was simply quoting what Boltzmann actually DID (in the paper [B96]).

8) Request for evidence on the acceptance of Boltzmann's ideas on irreversibility and ergodicity: the evidence is my experience, in the sense I have explained in the added note, which I realize from the referee's comment that shold have been written before.

9) I do not understand the referee's complaint about the gaussian thermostat and on it being a "quantity". Nor his comment on abstractness.

10) The footnote 9 is NOT a matter of priority: the latter, if necessary, is clearly established by the publication dates, and not much has to be discussed about that. I could not resist "preaching" on the not uncommon habit of not quoting known but unpleasant references. Therefore I have no problems in taking away most of the footnote and leave it just in the form of a quotation of an "unpleasant paper", which in spite of that deserves quotation, because it is a correct and interesting paper.

11) I cannot omit the discussion on the consequences of the "Ehrenfests shift" on the meaning of ergodic. I think this is an important point of my paper (and by the way it does not take much space in it). I still remember various statistical mechanics courses I attended as a student and the aura of confusion that the misterious quasi ergodic hypothesis brought in.

12) Yes the discussion on Boltzmann views on discreteness might be confusing and conjectural: nevertheless it has been repeatedly considered of great interest (I quoted Dugas and Klein to support this) and I hoped to have said something to stress once more its interest and that something can be said about it, that perhaps had not yet been said.

13) The last section of the paper is, in my view, consistent both logically and technically with the rest of the paper. The example I discuss is quite real, in fact better than real as it hints at a possibility of prediction of the result of an experiment. I cannot develop the matter further as that is part of a series of papers that I am writing also in collaboration with other colleagues. I would like to stress that sec.4 is a conceptual program, that

I think is quite detailed as such. I thought it would be nice to explain the program and its connection with the roots of the statistcial theory.
End of comments

To the editor:

I can see that the paper is not on the mainstream of the history of Physics. This is only partly due to the attempt at showing how a certain way of looking at statistical mechanics is still very much alive, even though to non experts it looks very different from what we do today. There seem to be certain deep disagreements with the referee: however I just disagree with him and I do not see how I could improve that, besides the changes I made because of his remarks. Furthermore not being on the mainstream is, in my view, part of its interest.
...
One thing I would really dislike is cutting sec. 4 out. My all point is that it is intimately related with the ergodic hypothesis of Boltzmann.
G.G.

Teoria Ergodica.

Voce per l'Enciclopedia del Novecento
dell'Istituto della Enciclopedia Italiana
Autore: Giovanni Gallavotti

Roma, dicembre 1986

§1 Le origini.

È comune opinione che la teoria ergodica nasca dall'idea di Boltzmann che sia possibile dare una giustificazione microscopica della legge di Boltzmann–Gibbs sul "metodo" di calcolo dei valori medi nel tempo di grandezze osservabili su un sistema di $N \approx 10^{20}$ particelle senza tuttavia risolvere le equazioni del moto. Peró, sebbene sia vero che la teoria ergodica nasca dalle idee di Boltzmann, sarebbe fuorviante credere che emerga dalla sua teoria della legge di Boltzman–Gibbs e ridurrebbe il ruolo di Boltzmann quasi a casuale ispiratore di problemi matematici irrilevanti per la Fisica. E' perció utile un breve esame del pensiero di Boltzmann al fine, non tanto di puntualizzare il suo ruolo, quanto di introdurre molti dei problemi della teoria ergodica moderna (spesso accusata di astrattezza matematica) in forma direttamente connessa a questioni fisiche rilevanti.

Il "metodo" di calcolo é semplice.

Denotiamo simbolicamente $\underline{v}, \underline{q}$ le coordinate di velocità e di posizione delle N particelle di massa m (per semplicità identiche) e denotiamo $S_t(\underline{v}, \underline{q})$ i valori di queste coordinate dopo un tempo t durante il quale il sistema evolve con un moto descritto dalle leggi della dinamica.

Si suppone che le particelle interagiscano a mezzo di una forza con energia potenziale uguale alla somma delle energie potenziali delle forze reciproche fra le varie coppie di particelle. Inoltre si suppone che queste forze abbiano "corta" portata, ossia agiscano su distanze dell'ordine di grandezza delle distanze interatomiche nei cristalli. Infine il sistema si immagina che il sistema sia racchiuso in un contenitore $\underline{\omega}$ a pareti perfette.

Allora il valore medio di una grandezza osservabile, ossia di una funzione $f(\underline{v}, \underline{q})$ dello stato microsopico del sistema (come ad esempio il numero di particelle in un cubetto unitario, ovvero la somma delle loro energie cinetiche $\frac{1}{2}\sum_{i=1}^{N} m\underline{v}_i^2$, o ancora la energia potenziale totale *etc.*), é:

$$\overline{f} = \lim_{T\to\infty} \frac{1}{T}\int_0^T dt\, f\big(S_t(\underline{v}, \underline{q})\big) = \begin{array}{l}\text{valore medio di } f \text{ nelle}\\ \text{configurazioni di energia}\\ \text{uguale al valore iniziale}\end{array} \tag{1.1}$$

Questa relazione vuol dire che se si fissa un intervallo di tempo T molto grande rispetto al "tempo microscopico" τ sulla cui scala sono percepibili i movimenti delle particelle e se il valore dell' osservabile f agli istanti $i\tau$, $i = 0, 1, \ldots$ si denota $f_i \equiv f(S_{i\tau}(\underline{v}, \underline{q}))$ allora la quantità:

$$\frac{f_0 + f_1 + \ldots + f_M}{M} \qquad \text{ove } M = \frac{T}{\tau} \tag{1.2}$$

che per essere calcolata necessiterebbe dei valori f_i ossia dei punti $S_{i\tau}(\underline{v}, \underline{q})$ e, cioé, delle soluzioni delle equazioni differenziali del moto, é secondo Boltzmann anche calcolabile in altro modo.

Si considerino tutti gli stati microscopici $(\underline{v}, \underline{q})$ e si immagini di suddividere lo spazio a $6N$ dimensioni, in cui tali variabili variano, in piccoli elementi di volume parallelepipedi, che chiameremo "cellette", in cui ciascuna componente delle velocità varia di δv e ciascuna componente delle posizioni varia di δq: ad ognuna di queste cellette si assoceranno i valori T, V e $E = T + V$ delle energie cinetica, potenziale e totale delle particelle. Si deve immaginare che le cellette abbiano le piú piccole dimensioni compatibili con la precisione con cui possiamo immaginare di eseguire le misure delle coordinate in questione, almeno in linea di principio: quindi ogni celletta ci fornisce *tutta* la possibile informazione sullo stato microscopico del sistema.

Consideriamo ora tutte le cellette con energia uguale al valore E dell' energia iniziale del nostro sistema. Poiché il sistema é racchiuso in un contenitore Ω finito (anche se macroscopico) il numero $\mathcal{N}$ delle cellette che ne rappresentano gli stati possibili con energia totale E é molto grande ma finito. Potremo così numerarle $\Delta_1, \ldots, \Delta_{\mathcal{N}}$, e la prescrizione alternativa, di Boltzmann, per il calcolo di (1.1) si formula così:

$$\lim_{M\to\infty} \frac{f_1 + \ldots + f_M}{M} = \frac{\sum_{i=1}^{\mathcal{N}} f(\Delta_i)}{\mathcal{N}} \tag{1.3}$$

ove $f(\Delta_i)$ é il valore che la grandezza osservabile in questione assume sulla $i - ma$ celletta della famiglia di cellette di uguale energia.

La (1.3) é in sostanza la *ipotesi ergodica* nella sua formulazione originale: bisogna infatti tener conto che Boltzmann realmente pensava alle variabili di posizione e velocità come variabili in cellette di dimensioni così piccole da non poter fisicamente distinguere i punti al loro interno e su questa rappresentazione costruì tutta la sua intuizione sulla meccanica microscopica e le sue teorie su di essa.

Non é compito di questo articolo entrare nella discussione del perché la (1.3) sia un' idea così rivoluzionaria, sulla quale sia possibile fondare una spiegazione della Termodinamica macroscopica in termini del modello atomico della materia: basterà osservare che, sebbene il membro di destra appaia (e realmente sia) difficile da calcolare, tuttavia sostituisce il problema del calcolo di una media temporale con quello, in un certo senso geometrico-combinatorio, di calcolare il secondo membro della (1.3) in cui la dinamica appare non già attraverso la soluzione delle corrispondenti equazioni differenziali ma semplicemente attraverso la funzione energia totale (e, cioé, per selezionare le $\mathcal{N}$ cellette con energia uguale a quella del dato iniziale, fra tutte le infinite cellette in cui si puó immaginare idealmente diviso lo spazio "delle fasi" $\underline{v}, \underline{q}$).

Boltzmann si rese subito conto che la ipotesi (1.3), che aveva già in precedenza mostrato implicare una spiegazione microscopica della Termodinamica (in un classico lavoro del 1884), avrebbe potuto avere la seguente interpretazione dinamica.

La dinamica newtoniana conserva l'energia e dunque, se le cellette sono sufficientemente piccole, la dinamica é pensabile, quando osservata ad intervalli di tempo τ dell'ordine di grandezza dei tempi microscopici sui quali le varie grandezze atomiche variano sensibilmente, come una trasformazione che ad ogni celletta Δ associa una celletta Δ' di uguale energia che é quella in cui Δ evolve (una sola, con buona approssimazione se τ é piccolo, in virtú di proprietà fondamentali della dinamica newtoniana in base alle quali il volume della celletta resta invariato nel tempo (teorema di Liouville) e inoltre (unicità della dinamica) due diverse cellette non possono essere trasformate nella stessa).

In altre parole la dinamica agisce sulle cellette dello spazio delle fasi di assegnata energia *permutandole* fra loro. E allora l' ipotesi ergodica é che questa permutazione é molto semplice e consiste in *un solo ciclo*: una data celletta Δ_1 evolve successivamente in cellette diverse in modo che dopo un numero $\mathcal{N}$ di passi pari al numero totale delle cellette ha visitato tutte quelle di uguale energia e ritorna allo stato iniziale.

Le conseguenze di una ipotesi così audace sono molteplici: la prima é proprio la uguaglianza fra le medie temporali di una qualunque grandezza osservabile e le sue medie (c.f.r. (1.3)) sulla superficie degli stati di uguale energia, nello spazio delle fasi. La conoscenza della dinamica porrebbe un ordine ciclico naturale fra le cellette di uguale energia: ma tale ordine é irrilevante ai fini dl calcolo della media in (1.3). Inoltre segue immediatamente che, in questa rappresentazione, tutti i moti sono periodici (e di periodo $\mathcal{N}\tau$): é il fenomeno della "ricorrenza".

Peró ci sono anche varie conseguenze negative che subito sembrarono prevalere su quelle positive e valsero a Boltzmann le critiche quasi unanimi dei suoi contemporanei, aggravate dalla fondamentale incomprensione della sua opera da parte degli Ehrenfest che ne furono i primi divulgatori (con il loro articolo del 1911 sul Encyklopädiae der Mathematischen Wissenshaften [1]).

Una ovvia conseguenza negativa é che tutti i moti sarebbero sì periodici, ma occorrerebbe ovviamente attendere tempi enormemente lunghi prima che il sistema possa ritornare sui suoi passi.

Boltzmannin un altro lavoro (1896), tanto famoso quanto in realtà poco noto dal punto di vista tecnico, dá una stima sulla lunghezza di questo ciclo nel caso di un piccolo campione di gas: e risulta dell'ordine di $10^{10^{19}}$ anni,

[1] tale incomprensione la si puó apprezzare forse già dall'etimologia errata che attribuiscono alla parola ergodico secondo la quale deriverebbe da ἔργον(energia) e ὁδός(via), mentre nel lavoro del 1884 é chiaro che deriva da ἔργον(energia) e da εἶδος(aspetto) ad abbreviazione di "ergomonodo" ove monodo (da μόνος(unico) e εἶδος) sta a significare quello che oggi é chiamato un insieme statistico stazionario.

ossia di gran lunga superiore alla piú lunga scala di tempo immaginabile (inclusa l' età dell' universo che é di soli $\approx 10^{10}$ anni!).

Come fu, assai presto se non immediatamente, chiaro a Boltzmann l' ipotesi ergodicanon puó essere di per sé presa a giustificazione della (1.3). E invero Boltzmann usó questa stima non per giustificare la (1.3) ma per ribattere alle critiche di coloro che si opponevano alla equazione di Boltzmann in quanto quest'ultima prediceva irreversibilità in apparente contrasto con la dinamica microscopica reversibile e ciclica.

La giustificazione che Boltzmann realmente adottó della (1.3) in quanto formula su cui fondare la teoria microscopica della termodinamica e rendere compatibili l'ipotesi ergodica e le relative scale di tempo assurdamente lunghe, fu che le grandezze fisicheosservabili *f rilevanti per la termodinamica* fossero assai poche e che avessero la proprietà fondamentale di essere praticamente costanti, e quindi uguali al loro valore medio, a destra della (2.3), sulla superficie di data energia: per cui, dopo poco tempo il sistema avrebbe visitato talmente tante cellette da aver fatto acquisire, almeno alle osservabili in questione, un valore medio temporale pari al valore medio teorico su tempi di lunghezza infinita (dato appunto dalla ipotesi ergodica).

Questa idea era naturale sbocco del pensiero di Boltzmann che già nel lavoro del 1884 faceva vedere per quale ragione le grandezze fisiche rilevanti per la termodinamica (e cioé quelle i cui valori medi fornivano i valori della pressione, temperatura, energia interna, densità, *etc*) avessero valore praticamente costante. E pochi anni prima (1872) aveva fornito, tramite l' equazione di Boltzmann, una spiegazione teorica del meccanismo per cui il tempo necessario perché una media temporale venga a coincidere con la media sullo spazio delle fasi risulta essere di ordine di grandezza osservabile e accettabile (nulla avendo a che vedere con i tempi superastronomici di ricorrenza).

In conclusione l' ipotesi ergodica, già nel pensiero di Boltzmann, viene a perdere il ruolo di fondamento della termodinamica e della meccanica statistica che é preso da teorie piú dettagliate e in un certo senso piú difficili (quali l' equazione di Boltzmann) nelle quali viene usata solo come condizione di consistenza (cioé i valori medi temporali raggiunti in tempi relativamente brevi possono restare costanti indefinitamente se il loro valore é appunto la media sulla superficie di energia costante, sulla quale il sistema si evolve ergodicamente).

È peró interessante osservare che, in questa luce, l' ipotesi ergodica non é neppure veramente necessaria: infatti per le applicazioni sarebbe sufficiente la validità della (1.3) solo per le poche grandezze f di interesse (e non per tutte le possibili funzioni sullo spazio delle fasi) e per questo basterebbe l'ipotesi che, nella sua evoluzione, la celletta visitasse in grande prevalenza

le cellette sulle quali le osservabili di interesse hanno valore pressocché costante (e che sono la maggioranza delle cellette di data energia).

Tutti questi temi si ritrovano nella problematica moderna sulla teoria ergodica che si indirizza non solo alla questione se, in casi concreti, sia valida la (1.3) ma anche alle questioni forse piú interessanti che riguardano la validità parziale della (1.3) (cioé ristretta a poche selezionate f) e le stime delle varie scale di tempo coinvolte nel problema della velocità con cui il membro di sinistra della (1.3) tende al valore del membro di destra. (che è il prblema dell'avvicinamento all'equilibrio).

§2 *La struttura astratta della teoria ergodica moderna.*

Senza aver presente la succinta analisi sull'opera di Boltzmann appena esposta sarebbe difficile apprezzare l'unità concettuale dietro la classe di problemi che vengono riconosciuti piú o meno universalmente come parte della teoria ergodica e dei quali ora esamineremo una selezione significativa.

Per illustrare l'evoluzione della teoria ergodica da Boltzmann ad oggi conviene avere in mente sempre l'esempio di un " sistema dinamico finito" accato a quello di "sistema dinamico continuo" che é il vero oggetto della teoria moderna.

Entrambi questi concetti si sono già incontrati nella discussione del §1:
a) un sistema dinamico finito é una coppia $(\mathcal{M}_o, T_o)$ in cui $\mathcal{M}_o$ é un insieme con un numero finito di elementi, che denoteremo $e_1, e_2, \ldots, e_{\mathcal{N}}$ e T_o é una trasformazione degli elementi di $\mathcal{M}$. L'esempio del §1 fornisce il caso in cui $\mathcal{M}_o$ é lo spazio delle cellette con data energia e T_o é lo loro permutazione generata dalla dinamica.
b) un sistema dinamico continuo é invece una coppia $(\mathcal{M}, T)$ in cui $\mathcal{M}$ é una superficie in uno spazio euclideo e T é una trasformazione di $\mathcal{M}$ in se stessa. Converrà sempre immaginare che $\mathcal{M}$ sia liscia (cioé priva di singolarità e di bordo) e che T sia una trasformazione regolare (cioé dotata di infinite derivate). L'esempio del §1 fornisce il caso in cui $\mathcal{M}$ é la superficie di energia costante nello spazio delle fasi (a $6N$ dimensioni) in cui le pareti del contenitore sono modellate da un potenziale ripidissimo ma regolare e nullo eccetto che nei pressi della frontiera (si evita così che $\mathcal{M}$ abbia bordo). La T é in questo caso la trasformazione $S_\tau(\underline{v}, \underline{q})$ associata alla dinamica newtoniana del sistema su un tempo τ dell'ordine dei tempi microscopici sui quali la dinamica é percepibile.

Nel seguito, come già nel paragrafo precedente, converrà pensare che $(\mathcal{M}_o, T_o)$ sia un modello discreto di $(\mathcal{M}, T)$ ottenuto suddividendo $\mathcal{M}$ in cellette Δ molto piccole, ma con dimensioni non nulle per modo che in $\mathcal{M}$

ce n'é solo un numero finito e immaginando che T sia approssimata da una trasformazione T_o di $\mathcal{M}_o$.

E' anche interessante osservare che uno studio numerico di un sistema dinamico ha proprio come primo passo quello di sostituire $(\mathcal{M}, T)$ con un sistema dinamico finito $(\mathcal{M}_o, T_o)$ che é descritto nel programma del calcolo stesso e che é considerato una buona approssimazione di $(\mathcal{M}, T)$.

Il problema basilare nella teoria dei sistemi dinamici é il seguente: assegnato un dato iniziale $x \in \mathcal{M}$ come si puó descrivere il moto, cioé la successione $\{T^k x\}$, $k = 0, 1, \ldots$?

La prima nozione rilevante a questo fine é quella di "statistica" del moto con origine x. Si dice che il moto con dato iniziale x *ha una statistica ben definita* e descritta da una distribuzione di probabilità μ, su $\mathcal{M}$, se per ogni funzione f su $\mathcal{M}$, si ha:

$$\lim_{R\to\infty} \frac{1}{R} \sum_{k=0}^{R-1} f(T^k x) = \int_{\mathcal{M}} f(y)\mu(dy) \tag{2.1}$$

se, cioé, la media delle osservazioni del valore di f sul moto generato da x al variare del tempo puó essere calcolata immaginando che $T^k x$ spenda in una qualunque regione misurabile E una frazione del tempo totale trascorso pari alla probabilità $\mu(E)$ dell'insieme E rispetto alla distribuzione di probabilità μ [2].

Per apprezzare il significato di questa definizione, che non si deve confondere con la ipotesi ergodica con la quale verrà in seguito confrontata, é utile vedere cosa diviene nel caso in cui il sistema dinamico é finito e T_o é una permutazione di $\mathcal{M}_o$ (cioé una trasformazione invertibile).

In tal caso, poiché T_o é una permutazione degli $\mathcal{N}$ stati di $\mathcal{M}_o$, é chiaro che x dovrà appartenere ad un ciclo della permutazione che, necessariamente, consiste di x e delle sue successive immagini rispetto all'azione di T_o. Se dunque $x_1 \equiv x, x_2, \ldots, x_c$ é il ciclo in questione (con $x_2 = T_o x_1, \ldots, x_c = T_o x_{c-1}, x_1 = T_o x_c$), é chiaro che il moto di x é periodico con periodo c e quindi:

$$\lim_{R\to\infty} \frac{1}{R} \sum_{j=0}^{R-1} f(T_o^k x) = \frac{f(x_1) + \ldots + f(x_c)}{c} \tag{2.2}$$

e cioé x ha necessariamente statistica definita e la distribuzione di probabilità μ é la distribuzione che assegna probabilità $\frac{1}{c}$ a ciascuno dei c punti del ciclo di x e 0 agli altri (che ci sono realmente se $c < \mathcal{N}$, ossia se T_o é una permutazione a piú di un ciclo).

[2] quì si intenderà, come di solito, che la distribuzione di probabilità sia abbastanza ricca da potersi parlare di probabilità di tutti gli insiemi aperti, almeno (e cioé che tutti gli insiemi aperti siano misurabili).

Il caso in cui T_o non é invertibile (o "dissipativa", perché questo caso si verifica nei modelli di sistemi in cui c'é dissipazione) é facilmente riducibile al precedente: infatti poiché $\mathcal{M}_o$ contiene solo un numero finito di punti il moto di un qualunque dato iniziale deve diventare prima o poi periodico. Dunque $\mathcal{M}_o$ contiene un sottoinsieme $\mathcal{M}_{per}$ di punti il cui moto é periodico, e ogni altro punto dopo un numero finito di trasformazioni entra in $\mathcal{M}_{per}$: ovviamente la T_o ristretta all'insieme $\mathcal{M}_{per}$ é invertibile e pertanto ne é una permutazione. Gli altri punti di $\mathcal{M}_o$ rappresentano fenomeni transienti nel senso che le proprietà statistiche dei moti dipendono solo dall'insieme $\mathcal{C}$ dei cicli di T_o (come trasformazione di $\mathcal{M}_{per}$).

Questo esempio peró insegna molto di piú: i punti del sistema dinamico $(\mathcal{M}_o, T_o)$ si possono classificare in base alle statistiche che generano; e questo equivale a classificarli secondo i cicli di T_o poiché ogni statistica μ sarà necessariamente, come segue dalla sua definizione, una distribuzione di probabilità che dà probabilità uguale agli elementi di uno stesso ciclo e probabilità nulla agli altri.

Inoltre c'é anche corrispondenza fra i cicli di T_o, e certi "insiemi invarianti" di $\mathcal{M}_o$, ossia certi insiemi E di punti di $\mathcal{M}_o$ che sono trasformati in se stessi dalla trasformazione T_o: $T_oE = E$.

Precisamente si vede che ogni ciclo (e quindi ogni statistica possibile) puó essere pensato come un insieme invariante minimale, privo cioé di sottoinsiemi invarianti piú piccoli.

Infine risulta, sempre nell'esempio in discussione, che esistono molte distribuzioni di probabilità invarianti rispetto all'azione di T_o, cioé tali che $\mu(E) = \mu(T_o^{-1}E)$. Ogni statistica é naturalmente una di queste; e piú in generale ogni misura invariante μ deve essere una combinazione di statistiche:

$$\mu(E) = \sum_{\gamma \in \mathcal{C}} p_\gamma \mu_\gamma(E) \tag{2.3}$$

ove μ_γ é la statistica relativa al ciclo γ di T_o, e $\mathcal{C}$ denota l'insieme dei cicli di T_o e i coefficienti p_γ sono non negativi e tali che $\sum_\gamma p_\gamma = 1$ (ed hanno l'interpretazione di probabilità rispetto a μ dei vari cicli). Ogni distribuzione invariante su $\mathcal{M}_o$ si decompone in somma di misure invarianti minimali (cioé non ulteriormente decomponibili) che sono appunto le statistiche dei vari punti di $\mathcal{M}_o$.

Ne emerge una rappresentazione piuttosto semplice delle proprietà statistiche dei moti di $(\mathcal{M}_o, T_o)$, e ci si domanda quanto di essa resti passando all'esame di sistemi dinamici continui e se per questi si possano elaborare considerazioni analoghe.

L' analogia c'è ed é essenzialmente perfetta: la sua scoperta ha segnato la nascita della teoria ergodica come é intesa oggi. Cominciamo con il porre alcune definizioni ispirate dalle proprietà dei sistemi finiti, appena messe

in luce.

Dato un sistema dinamico $(\mathcal{M}, T)$ ed una distribuzione di probabilità μ su $\mathcal{M}$ si dirà che μ é una distribuzione invariante se per ogni insieme (misurabile) $E \subset \mathcal{M}$ si ha $\mu(E) = \mu(T^{-1}E)$.

Si dirà che una misura μ é indecomponibile, o "ergodica", o metricamente transitiva, se non é possibile scrivere $\mu = \alpha\mu_1 + (1-\alpha)\mu_2$ con $\mu_1 \neq \mu_2$, μ_1, μ_2 invarianti e differenti e $0 < \alpha < 1$.

Poiché la proprietà di ergodicità di μ é strettamente legata a T e non avrebbe senso dire che μ é ergodica senza riferirsi ad una trasformazione T, si dice anche equivalentemente che T "agisce ergodicamente" sulla distribuzione invariante μ.

E i primi risultati della teoria ergodica sono stati appunto i due teoremi che garantiscono i) l'esistenza della statistica associata ad ogni punto di $\mathcal{M}$, in analogia alla (2.2), e ii) la possibilità di una decomposizione "ergodica" di una arbitraria distribuzione di probabilità invariante in misure invarianti indecomponibili (e quindi ergodiche), in analogia alla (2.3).

Piú precisamente si trova che ("teorema della decomposizione ergodica") data una distribuzione invariante μ é possibile rappresentarla come combinazione lineare di misure ergodiche in modo essenzialmente unico: peró, poiché in generale un sistema dinamico puó ammettere infinità anche non numerabili di misure invarianti indecomponibili, questa rappresentazione é in realtà un integrale anziché una somma e non entriamo qui nei dettagli tecnici necessari per precisare questa nozione.

Inoltre si trova ("teorema ergodico") che tutti i punti di $\mathcal{M}$ hanno statistica ben definita ed ergodica, con la possibile eccezione di un insieme di punti che peró ha probabilità nulla rispetto a qualsiasi distribuzione di probabilità invariante.

Si vede dunque che nel passaggio dal finito al continuo non avviene nulla di veramente nuovo e sorprendente; e il solo reale problema sta nel porre le definizioni in modo che abbiano il significato che si vuole. Sebbene la dimostrazione delle affermazioni ora fatte richieda spesso varie pagine nei trattati di teoria dei sistemi dinamici, rimane il fatto che quanto detto finora ha carattere generale, quasi esclusivamente definitorio, e deve essere considerato come una parte formale *preliminare* all'attacco di problemi di interesse applicativo.

Un' ulteriore questione di carattere generale é la seguente. Cosa accade se il punto x che si studia viene scelto a caso con un' assegnata distribuzione di probabilità, non necessariamente invariante, $\overline{\mu}$? Questa è una situazione che si presenta spesso nelle applicazioni, e la esamineremo prima nel caso di un sistema dinamico finito $(\mathcal{M}_o, T_o)$.

Ogni punto x ha anche in questo caso una statistica ben definita e la questione naturale é se tale statistica dipenda o meno dal punto x scelto a

caso con distribuzione $\overline{\mu}$, o almeno se questo avviene con probabilità 1.

Se, come stiamo supponendo, il sistema dinamico é finito, la risposta é chiara: la statistica di x é indipendente da x con $\overline{\mu}$–probabilità 1 se e solo se $\bar{\mu}$ dà probabilità non nulla agli elementi di un solo ciclo. Se invece c'é probabilità non nulla di scegliere elementi x che appartengono a cicli diversi allora é chiaro che ci sarà probabilità non nulla di trovare elementi x a statistica diversa.

E' allora naturale estendere la nozione di ergodicità di una distribuzione di probabilità alle distribuzioni non invarianti dicendo che una distribuzione di probabilità $\bar{\mu}$ su $\mathcal{M}$ é ergodica per il sistema dinamico $(\mathcal{M}, T)$ se con $\bar{\mu}$–probabilità 1 la statistica di un punto x scelto a caso con distribuzione $\bar{\mu}$ non dipende da x, cioé:

$$\lim_{R\to\infty} \frac{1}{R} \sum_{j=0}^{R-1} f(T^j x) = \int_{\mathcal{M}} f(y)\mu(dy) \tag{2.4}$$

con μ indipendente da x con $\bar{\mu}$–probabilità 1.

Si puó allora riformulare l'ultimo problema e al tempo stesso generalizzare il primo problema posto in questo paragrafo così: data una misura di probabilità $\bar{\mu}$ su M determinare se é ergodica ed in tal caso determinare la misura μ che ne descrive la statistica. Si tenga peró presente che in generale non sarà ora vero che quasi tutti i punti hanno statistica definita: tuttavia ció sarà vero (per quanto detto sopra per le distribuzioni invarianti) se $\bar{\mu}$ dà probabilità nulla agli insiemi che hanno probabilità nulla rispetto a tutte le misure invarianti, che é una condizione assai spesso verificata a dispetto della sua apparente complessità.

Come esempio per rendere piú concreta l'ultima questione, e quindi tutte quelle poste in questo paragrafo, si consideri il caso in cui $\mathcal{M}$ sia lo spazio delle fasi di un punto materiale vincolato ad una superficie regolare e liscia supposto soggetto ad una forza conservativa data, ed avente una energia totale data. Se $(\underline{p}, \underline{q})$ denotano le coordinate di momento e posizione del punto e se $H(\underline{p}, \underline{q})$ é la funzione di Hamilton associata al sistema, allora dalla meccanica é noto (teorema di Liouville) che la misura $d\underline{p}d\underline{q}\delta(H(\underline{p}, \underline{q}) - E)$ (ossia la misura uniforme sulla superficie di energia E nello spazio delle fasi) é invariante per la evoluzione newtoniana (descritta a mezzo delle equazioni di Hamilton con hamiltoniana H).

Ci si chiede se scegliendo a caso un punto, sulla supeficie dei dati aventi energia E, con distribuzione proporzionale alla misura di Liouville (ossia con distibuzione equiprobabile sulla superficie di energia E) si trova un punto a statistica definita e indipendente dal campione stesso prescelto. (Se la scelta a caso avvenisse invece con una distribuzione che non é invariante ma é, tuttavia, espressa da una densità non uniforme sulla superficie di

energia costante ci si troverebbe nel contesto piú generale della ergodicità di distribuzioni non invarianti).

Questo é un problema che di solito vien chiamato il problema della ergodicità del sistema dinamico hamiltoniano in esame. La affermazione (1.1), cioé l' ipotesi ergodica di B Boltzmann, puó essere tradotta nella affermazione che la misura di Liouville é ergodica rispetto al sistema dinamico in questione.

La risposta alla prima domanda é, per il teorema ergodico, affermativa: quasi tutti i punti x hanno una statistica ben definita. Il secondo problema (che é appunto l'analogo della (1.1)), e cioé se quasi tutti i punti abbiano la stessa statistica é invece assai difficile e ne é nota la soluzione solo per pochi casi specialissimi.

Ma la teoria ergodica si occupa principalmente di casi specialissimi e dopo aver posto in evidenza nel prossimo paragrafo ancora alcuni altri problemi generali che nascono spontaneamente dalle osservazioni di Boltzmann sulle scale di tempo dell' evoluzione del sistema, dedicheremo il resto di questo articolo all'esame di problemi particolari.

§3 *Problemi sulle scale di tempo.*

Abbiamo visto nel §2 come un sistema dinamico $(\mathcal{M}, T)$ sia tale che i suoi punti possano essere classificati in termini della statistica che generano, ossia della distribuzione di probabilità μ che generano via la relazione:

$$\lim_{R\to\infty} \frac{1}{R} \sum_{i=1}^{R-1} f(T^i x) = \int_{\mathcal{M}} f(y)\mu(dy) \tag{3.1}$$

E' un fatto naturale (e non proprio banale, anche se ovvio nel caso dei sistemi dinamici finiti) che i punti di $\mathcal{M}$ hanno tutti una statistica ben definita salvo al piú un insieme di punti che ha probabilità nulla rispetto a qualunque distribuzione di probabilità invariante (e quindi nessuna rilevanza statistica, *c.f.r.* §2).

Peró, come mostrato in modo davvero convincente da Boltzmann (c.f.r §1), la vera questione rilevante per le applicazioni, una volta che il dato iniziale x si é scelto a caso rispetto ad una data distribuzione $\bar{\mu}$ di probabilità su $\mathcal{M}$, sarà i) se la μ definita dalla (3.1) effettivamente esiste (con $\bar{\mu}$–probabilità 1) e non dipende dalla scelta casuale di x (é il problema della ergodicità di $\bar{\mu}$) e, data anche una f particolare, ii) la stima di quanto tempo si deve aspettare (ossia quanto grande deve essere R in (3.1)) perché la differenza fra il membro di destra della (3.1) e la media a sinistra sia inferiore ad un errore prefissato (é il problema delle scale di tempo per il raggiungimento

del valore di equilibrio da parte di una data osservabile, c.f.r. §1).

In altre parole ci si domanda se quasi tutti i punti di $\mathcal{M}$ abbiano una statistica unica (se scelti a caso con una data distibuzione di probabilità $\bar{\mu}$) e, in caso affermativo, quanto tempo occorra attendere perché questa statistica si manifesti, eseguendo osservazioni di una data osservabile f.

A proposito della seconda questione é rilevante un'altra nozione che conviene introdurre prima di procedere oltre. E' la nozione di "mescolamento" che é interessante e in certa misura sorprendente e delicata (a dispetto della sua semplicità formale) perché ha un semplice contenuto intuitivo e, ciononostante, é un concetto intrinsecamente continuo non avendo senso introdurlo per i sistemi dinamici finiti.

Si dice che un sistema dinamico $(\mathcal{M}, T)$ é mescolante per una distribuzione invariante di probabilità se per ogni coppia di insiemi E, F si ha:

$$\lim_{k\to\infty} \mu(E \cap T^k F) = \mu(E)\mu(F) \tag{3.2}$$

a volte si dice anche, equivalentemente, che T "mescola" la distibuzione invariante μ.

Il motivo del nome é chiaro: se T mescola μ e se k é grande l'insieme $T^k F$ sarà talmente uniformemente sparpagliato su $\mathcal{M}$ (immaginando di eseguire le misure del volume a mezzo della misura μ) che la parte di esso che é dentro un dato insieme E é pari alla parte di tutto lo spazio che é coperta da $T^k F$ stesso (sempre misurando gli insiemi con la μ).

L'interesse della nozione per quel che concerne il problema delle scale di tempo proviene dalla osservazione che, a volte, in luogo della velocità di convergenza del limite (3.1) risulta interessante quella della espressione che si ottiene integrando ambo i membri di (3.1) rispetto alla misura $g(x)\mu(dx)$, con una certa scelta di g, ottenendo:

$$\lim_{R\to\infty} \frac{1}{R} \sum_{i=0}^{R-1} \int_{\mathcal{M}} f(T^i x) g(x) \mu(dx) = \big(\int_{\mathcal{M}} f(y)\mu(dy)\big)\big(\int_{\mathcal{M}} g(x)\mu(dx)\big) \tag{3.3}$$

e la (3.2), implicando la proprietà piú forte della (3.3):

$$\lim_{i\to\infty} \int_{\mathcal{M}} f(T^i x) g(x) \mu(dx) = \big(\int_{\mathcal{M}} f(y)\mu(dy)\big)\big(\int_{\mathcal{M}} g(x)\mu(dx)\big) \tag{3.4}$$

ne puó divenire uno strumento di analisi.

È anche chiaro che un sistema dinamico finito non banale non puó mescolare alcuna misura invariante μ: questo perché la funzione $\mu(\{x\}\cap\{T^k y\})=$ (probabilità che y divenga x dopo un tempo k) é una funzione periodica in k e quindi, se non costante, non puó aver limite per $k \to \infty$ (salvo nel caso banale in cui $\mathcal{M}_o$ contiene un solo punto con misura non nulla e quindi non

si puó scegliere $x \neq y$). Inoltre si puó vedere che la nozione di mescolamento implica quella di ergodicità.

Nei prossimi paragrafi ci occuperemo di vari casi in cui si puó dare una risposta alle questioni poste in questi primi tre paragrafi e a questioni connesse: si tratterà sempre di casi particolari perché, come ci si puó immaginare, esistono ben pochi risultati generali. Purtroppo i casi che possono essere studiati in qualche dettaglio riguardano problemi che, pur interessanti e notevoli per questioni di analisi o geometria, assai di rado riguardano sistemi dinamici di reale interesse per la Fisica (anche fuori dal campo della teoria dei gas e della meccanica statistica).

§4 *Alcuni esempi di problemi della teoria ergodica.*

Questo paragrafo é dedicato all'elenco di una famiglia di problemi che sono esempi interessanti in cui la teoria ergodica é rilevante,

Esempio 1) *Semplici modelli di evoluzione.*

Si consideri una trasformazione $x \to Tx$ definita per $x \in \mathcal{M} \equiv [0,1]$. Il sistema dinamico $(\mathcal{M}, T)$ é un modello di evoluzione. Ben noto é il caso particolare $x \to \alpha x(1-x)$, con $\alpha \in [0,4]$ che é noto come "modello logistico" per l'evoluzione di una popolazione in cui x rappresenta la frazione del massimo valore, *a priori* possibile, effettivamente presente ad un dato istante.

Una tipica domanda che ci si puó porre é se, partendo da un dato iniziale scelto a caso con distribuzione $\mu(dx) = \rho(x)dx$, descritta da una densità ρ, il sistema evolva con statistica unica e quali siano le caratteristiche di questa statistica.

La risposta dipende evidentemente dalla struttura della funzione f. Per formarci un'idea della complessità che questa risposta puó avere consideriamo una famiglia T_α ad un parametro di sistemi dinamici del tipo considerato. Ed esaminiamo, al crescere di α fra 0 e α_o, il comportamento statistico. Occorre, per non restare troppo nelle generalità, ancora qualche ipotesi sulla forma delle funzioni T_α: é possibile formulare ipotesi molto generali che garantiscano che i fenomeni che si manifestano al variare di α siano qualitativamente e in molti sensi anche quantitativamente quelli che si manifestano nel caso della famiglia di trasformazioni logistiche appena introdotta. Per questo motivo, e per non appesantire l'analisi con condizioni tecniche, esporremo i risultati nel caso speciale delle trasformazioni logistiche.

Si trova che per valori di α piccoli la statistica asintotica é banale nel

senso che tutti i punti dell'intervallo hanno la stessa statistica la quale é semplicemente "probabilità 1 di trovare il sistema in 0": ció é dovuto al fatto che 0 é un "attrattore" ($\lim_{k\to\infty} T_\alpha^k x = 0$, per ogni x). Ovviamente la distribuzione di probabilità in questione é l'unica invariante ed é ergodica come pure qualsiasi altra distribuzione di probabilità su $[0, 1]$.

"Nulla accade" se α é piccolo. Ma se α cresce (nel caso in esame se diviene maggiore di 1) il sistema diviene assai piú interessante: dapprima la statistica asintotica é descritta da un moto periodico di periodo inizialmente 2 poi 4 poi 8 e così via, raddoppiando. Si puó dire che la statistica dei moti é ancora molto semplice, perchè con probabilità 1 tutti i dati iniziali evolvano verso un moto periodico (e quindi la statistica é ergodica ed é semplicemente una distribuzione di probabilità che attribuisce uguale probabilità ad un numero finito di punti e zero agli altri). In questi casi, peró, questa non é l'unica statistica possibile: con probabilità zero rispetto alla distribuzione con cui abbiamo convenuto di scegliere il dato iniziale ($\bar{\mu} = \rho(x)dx$) il moto puó evolvere verso altre statistiche anche esse descritte da moti periodici,(peró instabili).

A questo punto ci si domanda se la statistica dei moti mantenga sempre la stessa natura periodica ovvero se divenga realmente piú complessa: é chiaro che quel che sarebbe davvero nuovo sarebbe l'apparizione di statistiche che non sono concentrate su orbite periodiche, ma coinvolgono insiemi che hanno la potenza del continuo. Questo sarebbe ad esempio il caso in cui la statistica fosse descritta da una distribuzione di probabilità dotata di una densità: ossia avente la forma $\sigma(x)dx$. In questo ultimo caso il moto dovrebbe essere asintoticamente assai piú complesso ed interessante e il valore di x non sarebbe piú predicibile, perché appunto affetto da fluttuazioni statistiche continue.

Naturalmente dal punto di vista pratico anche un moto periodico potrebbe simulare una distibuzione statistica continua se il suo periodo fosse molto grande e se i punti della sua orbita fossero abbastanza sparpagliati: tuttavia la differenza concettuale fra i due casi é chiara.

Nel caso del nostro sistema si verificano tutti i casi: al crescere di α l'orbita periodica stabile che descrive la statistica di quasi tutti i punti dell'intervallo $[0, 1]$ continua a raddoppiare di periodo mantenendo lo stesso periodo in intervalli della variabile α successivi e *sempre piú piccoli* e con dimensione che diviene piccola in modo che i rapporti fra due intervalli successivi tendano ad un limite $\delta = 4.6692...$ ("fenomeno" e "costante di Feigenbaum).

Quando α raggiunge il valore limite di circa 2.7.. cominciano ad apparire statistiche piú complesse: situazioni in cui il moto asintotico é periodico si alternano a situazioni in cui invece la statistica é assai piú complessa e descritta da distribuzioni con densità positiva.

Quando $\alpha \to 4$ é sempre piú frequente il caso che la statistica sia descritta da densità e quindi sia "impredicibile": si puó mostrare (Jacobson) che l'insieme dei valori di α cui corrisponde una statistica asintotica continua diviene molto denso intorno al valore $\alpha = 4$ (nel senso che la misura dei punti α in un intervallo $[4-\varepsilon, 4]$ in corrispondenza dei quali T_α ha una statistica continua, divisa per la misura ε dell'intervallo stesso tende a 1 per $\varepsilon \to 0$).

Dunque man mano che ci si avvicina ad $\alpha = 4$ il moto risulta sempre piú spesso "caotico". Inoltre in corrispondenza di molti valori di α la statistica possibile é unica ed é possibile stimare la velocità con cui il secondo membro di (3.4) tende a zero, quando f, g sono funzioni regolari su $[0,1]$, e tale stima é esponenziale (in i). Dunque il sistema dinamico agisce in modo mescolante sulla statistica che produce e dalla conoscenza della velocità di mescolamento e dalla teoria necessaria per ottenerla si possono trarre varie conseguenze interessanti come stime esplicte del tempo che é necessario attendere perché i limiti in (3.3),(3.1) vengano raggiunti entro una precisione prefissata.

Esempio 2) *Sistemi Hamiltoniani.*

Sia $\mathcal{S}$ una superficie regolare l–dimensionale (per fissare le idee, analitica), liscia e priva di attrito. Su di essa é posto un punto materiale soggetto all'azione di una forza conservativa di energia potenziale V analitica. Se denotiamo con $(\underline{p}, \underline{q})$ le coordinate canoniche di momento e posizione che variano nello spazio delle fasi $\mathcal{F}$, potremo considerare la hamiltoniana come una funzione $H(\underline{p}, \underline{q})$. Sia E fissato e $\mathcal{M} \equiv \mathcal{F}_E$ denoti l'insieme dei punti dello spazio delle fasi sui quali il valore della hamiltoniana (cioé dell'energia) é E. Supporremo che $\mathcal{F}_E$ sia una superficie finita regolare e priva di bordo.

Sia $\mu(d\underline{p}d\underline{q}) \propto \delta(H(\underline{p}, \underline{q}) - E) d\underline{p}d\underline{q}$ (ossia se μ é la misura uniforme sulla superficie di energia E), e sia $T \equiv S_\tau(\underline{v}, \underline{q})$, ove $\tau > 0$ é un tempo di osservazione prefissato, denota l'evoluzione hamiltoniana. Allora si ha che $(\mathcal{M}, T)$ é un sistema dinamico e μ é una misura invariante (detta misura di Liouville).

Il caso studiato da Boltzmann a proposito della meccanica statistica (c.f.r. §1) é un caso speciale di questo (in cui la superficie é R^{6N} *etc.*; e la misura di Liouville prende di solito il nome di "insieme statistico microcanonico").

Ci domandiamo ora se T agisca ergodicamente o addirittura in modo mescolante su μ: ma ben poco é noto in generale.

Consideriamo il caso in cui il sistema in questione é integrabile per quadrature in un intorno di $\mathcal{F}_E$: si puó mostare che la misura di Liouville μ non é ergodica.

Infatti, per ipotesi, si possono trovare coordinate canoniche angolo–azione

$(\underline{A}, \underline{\varphi})$ e un intorno di $\mathcal{F}_E$, che in queste coordinate ha la forma $V \times T^l$ con $V \subset R^l$ e T^l=(toro a l dimensioni), in modo che in queste coordinate la hamiltoniana assuma la forma $h_o(\underline{A})$, con h_o opportuna.

In queste coordinate si vede che il moto é tale che le $\underline{A}$ non variano nel tempo, mentre le variabili angolari ruotano con velocità costante, ciascuna sul suo cerchio: $\varphi_i \to \varphi_i + \omega_i(\underline{A})t,\ i = 1, 2, \ldots, l$, con $\omega_i(\underline{A}) = \frac{\partial h_o}{\partial A_i}$: "moto quasi periodico" con pulsazioni $(\omega_1, \ldots, \omega_l) = \underline{\omega}(\underline{A})$.

Dunque se $T = S_\tau(\underline{v}, \underline{q})$, ove τ é un intervallo di tempo usato per ritmare le osservazioni, il sistema $(\mathcal{M}, T)$ ammette l integrali primi, le funzioni $\underline{A}$, e così la misura di Liouville non puó essere ergodica rispetto all'azione della dinamica in esame.

E' questa infatti una proprietà generale: la ergodicità é impossibile qualora il sistema abbia integrali primi non banali. Il perché lo si capisce subito osservando che se A é una costante del moto allora la sua media temporale coincide con il suo valore iniziale e quindi se A non é costante non puó esssere uguale per tutti i punti, come invece dovrebbe essere se il sistema fosse ergodico.

Tuttavia si possono associare ai sistemi integrabili altri sistemi dinamici: precisamente quelli in cui $\mathcal{F}_E$ vien sostituito da insiemi invarianti piú piccoli (che ci devono essere se il sistema non é ergodico): ad esempio gli insiemi che nelle coordinate angolo–azione hanno la forma $\mathcal{M} = \{\underline{A}\} \times T^l$. Ed é naturale considerare su di essi la distribuzione di probabilità invariante μ' che é omogenea sul toro: $\mu'(d\underline{\varphi}) = d\underline{\varphi}/2\pi^l$.

In questo caso si vede che, per il moto su $\mathcal{M}$, la distribuzione μ' é ergodica per quei valori di $\underline{A}$ per cui le componenti di $\underline{\omega}(\underline{A})$ sono "reciprocamente irrazionali" nel senso che $\omega_1 n_1 + \omega_2 n_2 + \ldots + \omega_l n_l + n_{l+1} 2\pi/\tau \neq 0$ per qualunque scelta di l numeri interi $n_1, n_2, \ldots, n_{l+1}$ non tutti nulli. Tuttavia, si puó vedere che T non agisce in modo mescolante sulla misura invariante μ': é questo uno dei piú semplici e interessanti esempi in cui un sistema dinamico agisce in modo non mescolante su una distribuzione invariante.

Questo mette anche in luce la connessione della teoria ergodica con questioni della teoria dei numeri (via l' analisi della irrazionalità reciproca delle pulsazioni $\underline{\omega}$). Si potrebbe vedere come il problema dello studio delle scale di tempo, che in questo caso puó essere risolta in forma abbastanza soddisfacente, sia (in questo caso) collegato intimamente alla teoria delle frazioni continue, (v.).

Esempio 3) *Ancora sui sistemi hamiltoniani.*

I sistemi hamiltoniani sono una sorgente inesauribile di problemi per la teoria ergodica. Ne menzioniamo quì alcuni altri.

Si consideri un sistema meccanico che é vicino ad un sistema integrabile

nel senso che i suoi moti sono descritti, nelle coordinate angolo–azione di un opportuno sistema integrabile, da una hamiltoniana che ha la forma $H = h_o(\underline{A}) + \varepsilon v(\underline{A}, \underline{\varphi})$ con ε piccolo, ossia che dipende poco dalle variabili di angolo $\underline{\varphi}$.

Si puó dimostrare, e la dimostrazione é stata uno dei maggiori successi della meccanica teorica, che per ε abbastanza piccolo il sistema dinamico ottenuto, come negli esempi del capoverso 2), considerando i moti di questo sistema che si svolgono ad energia fissa non agisce in modo ergodico sulla misura di Liouville, esattamente come nel caso precedente ($\varepsilon = 0$). Infatti si possono costruire opportune funzioni che sono costanti del moto per la dinamica in questione ma che non sono costanti sulle superfici $H = E$.

La difficoltà della loro costruzione sta nel fatto che questi integrali primi sono funzioni non regolari delle coordinate: il metodo per la loro costruzione (dovuto a Kolmogorov) é molto interessante e il risultato ebbe come primo effetto di sfatare una diffusa credenza che, per la interpretazione impropria di un teorema di Poincaré, voleva che fosse sufficiente una piccola perturbazione per distruggere tutti gli integrali primi di un sistema integrabile e fare in modo che la dinamica agisse ergodicamente sulla distribuzione di Liouville.

Ma é possibile anche trovare sistemi hamiltoniani in cui la trasformazione T, cioé la evoluzione su un tempo τ, agisce ergodicamente sulla distribuzione di Liouville. Basta "allontanarsi" dai sistemi integrabili per trovare esempi intressanti.

L'esempio fondamentale (E. Hopf) é quello del moto libero di un punto materiale su una superficie 2–dimensionale a curvatura negativa costante. Questo esempio, importante per le sue connesssioni con la geometria non euclidea, é tale che non solo la misura di Liouville é ergodica rispetto a T ma é addirittura mescolante e, anzi, é possibile stimare la velocità di mescolamento (c.f.r. (3.4)) come esponenziale per tutte le funzioni f, g analitiche su $\mathcal{M}$. Le tecniche quì coinvolgono alcune delle proprietà ed idee fondamentali nella geometria non euclidea e questo fatto ha reso giustamente famoso questo caso come uno dei piú eleganti della teoria ergodica, sebbene non abbia finora ricevuto applicazioni fisiche dirette.

I risultati sulla ergodicità si estendono poi ai casi analoghi in cui la curvatura non é costante, pur restando negativa: ma in questi casi non sono note tecniche per la stima delle velocità di mescolamento né per la stima della velocità con cui vien raggiunto il limite (3.1).

Esempio 4) *Biliardi e diffusione.*

Si considera una regione quadrata ("tavolo") sulla quale si muove un punto materiale ("palla"): supporremo che la palla, qualora urti il bordo, riappaia

con la stessa velocità sul bordo opposto ("tavolo periodico"): questo sarà il caso quì esaminato lasciando da parte il caso, analogo, in cui la palla venga invece riflessa elasticamente ("tavolo riflettente").

Possiamo rappresentare lo stato della palla a mezzo delle due componenti della velocità (A_1, A_2) e a mezzo di due angoli (φ_1, φ_2) variabili fra 0 e 2π che individuano la posizione sul tavolo, assumendo che il suo lato sia lungo 2π. Si impongono in questo modo naturalmente le condizioni al bordo periodiche ed il tavolo viene ad essere pensato come un prodotto di due cerchi, ossia come un toro a 2 dimensioni.

Nel mezzo del tavolo supporremo che esista un campo di forze, in modo che l' energia totale del sistema abbia la forma:

$$\frac{1}{2}(A_1^2 + A_2^2) + V(\varphi_1, \varphi_2) \tag{4.1}$$

Il caso limite in cui V diviene infinita in certe regioni convesse $C_1, \ldots, C_n$ (pioli) e nulla fuori di esse é il caso trattato recentemente da Sinai, interpretando le collisioni con i bordi dei pioli come collisioni elastiche. Si considera la superficie $\mathcal{F}_E$ dei punti dello spazio delle fasi con energia E e su di essa la misura di Liouville: si deve peró osservare che il sistema dinamico che si ottiene non é un sistema dinamico nel senso fin quì utilizzato perché ora la superficie (3–dimensionale) $\mathcal{M} = \mathcal{F}_E$ ha bordo e la trasformazione T non é ovunque regolare presentando discontinuità nelle derivate in corrispondenza di punti che danno origine a traiettorie che hanno, nel corso della loro evoluzione fra $t = 0$ e $t = \tau$, collisioni con i pioli.

Ma le varie nozioni introdotte possono essere facilmente estese al caso di sistemi dinamici in cui T non é regolare e $\mathcal{M}$ ha bordo o, piú in generale é solo uno spazio metrico: cosí si puó dimostrare, se i pioli hanno ovunque curvatura non nulla, che nel caso limite in questione il sistema dinamico é *mescolante* per la misura di Liouville. E questo risultato, che ancora oggi non sembra ridotto ad un fatto veramente elementare, é uno dei piú profondi della teoria ergodica e la complessità della teoria necessaria per ottenerlo fà capire la intrinseca difficoltà dei problemi della teoria ergodica, quando si vogliano ottenere risultati fisicamente interessanti. Sembra inoltre che sia anche possibile riuscire a ottenere stime delle velocità di mescolamento; ma i risultati al momento appaiono ancora in uno stadio preliminare.

Il problema del biliardo periodico con ostacoli "rigidi" (cioé nel caso limite considerato da Sinai) nasce da un'interessante applicazione che é abbastanza semplice da poter essere esposta senza troppe nozioni ausiliarie.

Si immagini una pallina che si muove in un piano infinito ricoperto di ostacoli disposti periodicamente. Supponiamo di lanciare la pallina con velocità 1 a partire da una posizione e con una velocità casuali scelte con la

distribuzione di probabilità $\mu \propto d\underline{\varphi} d\vartheta$, ove $\underline{\varphi}$ denota la posizione della pallina in un quadrato (supposto, per comodità di esposizione, di lato 2π) del pavimento che costituisce il mosaico periodico degli ostacoli e ϑ é l'angolo che la velocità forma con un asse fisso (asse x).

Sia $r(t)$ la distanza dall'origine raggiunta dal dato iniziale $(\underline{\varphi}, \vartheta)$ in un tempo t (multiplo del tempo τ prefissato, al cui ritmo si eseguono le osservazioni del moto). Ci domandiamo quale é il valor medio sulle posizioni iniziali di $r(t)/t$ per $t \to \infty$. Il limite di questa quantità é il "coefficiente di diffusione" D attraverso gli ostacoli, se esiste.

Il problema del calcolo del coefficiente di diffusione puó essere impostato così: si scrive che $r(t)^2$ é il quadrato dell'integrale della velocità fra l'istante $t = 0$ e l'istante corrente:

$$\begin{aligned} \lim_{t\to\infty} \frac{r(t)^2}{t} &= \lim_{t\to\infty} \frac{1}{t} \big(\underline{\varphi} + \int_0^t \underline{A}(t')dt'\big)^2 = \\ &= \lim_{t\to\infty} \frac{1}{t} \int_0^t dt'dt'' \underline{A}(t') \cdot \underline{A}(t'') = \\ &= \lim_{t\to\infty} \frac{1}{t} \int_0^t dt'dt'' \cos(\vartheta(t') - \vartheta(t'')) \end{aligned} \tag{4.2}$$

ove $(\underline{\varphi}(t), \vartheta(t))$ denotano le coordinate di posizione (a meno di multipli di 2π =(lato del quadrato elementare del reticolo degli ostacoli)) e l'angolo formato dalla velocità $\underline{A}(t)$ con un asse fisso che corrispondono al dato iniziale $(\underline{\varphi}, \vartheta)$.

Pertanto mediando sulla scelta casuale (rispetto alla misura uniforme μ sui dati iniziali considerati) si trova dopo qualche passaggio e immaginando, per semplicità, che il reticolo degli ostacoli sia simmetrico per rotazioni di 90° in modo che non ci sia distinzioone fra gli assi cartesiani:

$$\begin{aligned} D =& 2 \int_0^\infty \int_{\mathcal{F}_1} \cos(\vartheta(t') - \vartheta)) dt' \mu(d\underline{\varphi} d\vartheta) = \\ =& 4 \int_0^\infty dt' \Big(\int_{\mathcal{F}_1} \cos(\vartheta(t')) \cos(\vartheta) \mu(d\underline{\varphi}, d\vartheta) \Big) \end{aligned} \tag{4.3}$$

purché la funzione integranda nell'ultimo integrale sia decrescente per $t \to \infty$, in modo che l'integrale sia sommabile.

Poiché tale sommabilità dipende dala velocità con cui avviene il "mescolamento" per la coppia di funzioni $f = g = \cos\vartheta$, ossia dalla velocità con cui l'integrale (3.3) tende a zero nel caso del sistema dinamico in esame, vediamo come la propietà di mescolamento possa intervenire in un problema concreto.

Si vede anche come interessi, per lo studio della diffusione in un reticolo regolare di ostacoli, solo la proprietà di mescolamento per una funzione

speciale f su $\mathcal{M}$, ($f = \cos\vartheta$), e *non* già per tutte le possibili funzioni. E' questa una proprietà che in linea di principio non ha molto a che vedere con l' ergodicità, né con il mescolamento per tutte le funzioni (c.f.r. §1 ove viene discussa una questione analoga per quel che riguarda le proprietà ergodiche minime che é sufficiente provare per dare una giustificazione microscopica della termodinamica).

Esempio 5)*La meccanica statistica dei sistemi infiniti.*

Non si deve pensare che la teoria ergodica abbia per oggetto solo problemi di evoluzione nel tempo, nel senso letterale della parola. Cosi', ad esempio, un altro dei campi piú fertili di esempi e applicazioni della teoria ergodica é la meccanica statistica.

Quì il ruolo della evoluzione nel tempo é giocato dalle traslazioni nello spazio. Il punto é che le nozioni di ergodicità e mescolamento si estendono facilmente al caso in cui la traslazione temporale é sostituita da un gruppo (o semigruppo) a piú generatori (nei casi finora considerati si puó immaginare che la trasformazione T e tutte le sue potenze formino un gruppo ad un generatore o, se T non é invertibile, un semigruppo).

Consideriamo, solo per semplicità, un sistema di particelle che possono occupare i punti di un reticolo quadrato Z^d di dimensione d (di solito $d = 3$). Lo spazio delle configurazioni che queste particelle possono assumere é $\mathcal{M} =$ {tutti i possibili sottoisiemi di Z^d}. Rappresenteremo una configurazione $\underline{\sigma}$ come una successione $\underline{\sigma} = \{\sigma_\xi\}_{\xi\in Z^d}$ indicizzata dai punti $\xi \in Z^d$ del reticolo e con $\sigma_\xi = \pm 1$. Il caso in cui le particelle possono occupare il continuo spaziale sarebbe analogo, ma con qualche difficoltà tecnica in piú.

Gli elementi $\underline{\sigma}$ dello spazio $\mathcal{M}$ rappresentano le configurazioni del sistema: si interpreta $\sigma_\xi = 1$ come significante che il sito $\xi \in Z^d$ é occupato da una particella, e $\sigma_\xi = -1$ come significante che il sito ξ é vuoto.

Su $\mathcal{M}$ agisce il gruppo delle traslazioni spaziali generato dalle d trasformazioni $\underline{T} \equiv (T_1, T_2, \ldots, T_d)$ di $\mathcal{M}$ in se; se $\underline{\sigma} \in \mathcal{M}$ allora, per $j = 1, 2, \ldots, d$:

$$T_j\underline{\sigma} = \underline{\sigma}' \qquad \text{con } \sigma'_\xi = \sigma_{\xi - e_j} \tag{4.5}$$

ove e_j é un vettore che congiunge due siti primi vicini su Z^d nella direzione j–ma. Le trasformazioni T_j "generano" il gruppo delle traslazioni spaziali delle configurazioni del sistema. E la coppia $(\mathcal{M}, \underline{T})$ definisce un "sistema dinamico" a d dimensioni, che generalizza la nozione data al §2 non solo perché ora il gruppo delle trasformazioni ha piú di un generatore ma anche perché lo spazio $\mathcal{M}$ non é una superficie.

Si puó ora immaginare di considerare quello che in meccanica statistica viene chiamato un "insieme statistico", cioé una distribuzione μ di proba-

bilità su $\mathcal{M}$ invariante rispetto alle trasformazioni di traslazione (é naturale, nello specificare quali siano gli insiemi misurabili, ossia quali siano gli insiemi di cui é possibile domandare quale sia la probabilità di realizzazione, richiedere che siano misurabili almeno tutti gli insiemi che possono essere definiti specificando le proprietà dei valori delle σ_i al variare di i in un insieme finito di indici: sono gli insiemi che giocano il ruolo degli insiemi aperti nella teoria della probabilità ordinaria; da un punto di vista fisico questi insiemi hanno la proprietà che l' appartenenza o meno ad essi, di un punto $\underline{\sigma}$ dato, può essere decisa semplicemente controllando l'aspetto di $\underline{\sigma}$ su un numero finito di siti). Una tale distribuzione di probabilità, per la sua proprietà di invarianza per traslazione, rappresenta uno "stato omogeneo" del sistema.

Le nozioni di ergodicità e mescolamento dell'azione di $\underline{T}$ sulla distribuzione invariante μ si danno in modo naturale:

$$\lim_{R\to\infty} \frac{1}{R^d} \sum_{j_1=0}^{R-1} \sum_{j_2=0}^{R-1} \cdots \sum_{j_d=0}^{R-1} f(T_1^{j_1} \dots T_d^{j_d} \underline{\sigma}) = \int_{\mathcal{M}} f(\underline{\sigma}')\mu(d\underline{\sigma}') \tag{4.6}$$

per μ–quasi tutte le scelte di $\underline{\sigma}$, definisce l'ergodicità; mentre il mescolamento é definito da:

$$\int_{\mathcal{M}} f(T_1^{j_1} \dots T_d^{j_d} \underline{\sigma})\, g(\underline{\sigma})\mu(d\underline{\sigma}) \xrightarrow[|\underline{j}|\to\infty]{} \int_{\mathcal{M}} f(\underline{\sigma})\mu(d\underline{\sigma}) \int_{\mathcal{M}} g(\underline{\sigma})\mu(d\underline{\sigma}) \tag{4.7}$$

per tutte le funzioni f, g.

Gli insiemi statistici che rappresentano,nella meccanica statistica, gli stati di equilibrio termodinamico si interpretano come "fasi pure" se sono ergodici rispetto all'azione delle traslazioni $\underline{T}$: e la decomposizione ergodica di uno stato di equilibrio termodinamico si interpreta come risoluzione di un miscuglio in fasi pure.

Concludiamo questo rapido riferimento alla meccanica statistica menzionando che la proprietà di mescolamento é rilevante, fra l'altro, per la discussione delle proprietà delle "fluttuazioni statistiche". Ad esempio, data una distribuzione μ invariante per traslazione, ci si puó domandare quale sia la struttura della variabile casuale $\nu_L(\underline{\sigma})$ definita da:

$$\nu_L(\underline{\sigma}) = \frac{\text{numero di particelle di } \underline{\sigma} \text{ in } \Omega_L - N_L}{\sqrt{\text{volume di } \Omega_L}} \tag{4.8}$$

ove Ω_L é un cubo di lato L e N_L é il valore medio del numero di particelle in Ω_L (sempre rispetto alla distribuzione μ).

In molti casi si arriva a dimostrare (ed é questo un altro tipico problema della teoria ergodica che potrebbe essere formulato molto in generale) che la variabile $\nu_L(\underline{\sigma})$ verifica una legge "normale":

$$\textit{probabilità}(\, \nu_L(\underline{\sigma}) \in (a,b)) \xrightarrow[L\to\infty]{} \int_a^b e^{-x^2/2\delta} \frac{dx}{\sqrt{2\pi\delta}} \tag{4.9}$$

con δ dato da:

$$\delta = \sum_{\underline{j} \in Z^d} \int_{\mathcal{M}} (\sigma_{\underline{j}} - \overline{\sigma})(\sigma_{\underline{O}} - \overline{\sigma}) \mu(d\underline{\sigma}) \tag{4.10}$$

se la serie a secondo membro é convergente (ossia se il mescolamento é abbastanza rapido), e se $\overline{\sigma}$ é il valore medio della variabile $\sigma_{\underline{O}}$ relativa al sito $\underline{j} = (\text{origine di } Z^d) \equiv \underline{O}$.

Si vede quí un'altra istanza in cui la rapidità di mescolamento appare come un ingrediente necessario per stabilire una proprietà di interesse diretto nelle applicazioni.

Si potrebbe continuare a lungo elencando problemi che si inquadrano nell'ambito della teoria ergodica. Ancora piú a lungo ci si potrebbe soffermare sulle tecniche impiegate per stabilire i vari risultati, tutte di grande interesse matematico e spesso di chiara interpretazione fisica. Non possiamo peró addentrarci in tali questioni e dedicheremo il prossimo paragrafo ad illustrare alcuni altri concetti fondamentali della teoria ergodica.

§5 *Un altro aspetto della teoria ergodica: entropia.*

Le questioni che riguardano le proprietà di ergodicità e mescolamento dell'azione di un sistema dinamico $(\mathcal{M}, T)$ su una distribuzione di probabilità μ, invariante, sono le prime e piú naturali questioni della teoria ergodica, ma, di gran lunga non ne esauriscono la classe di problemi.

La teoria delle fluttuazioni, alla quale abbiamo fugacemente accennato nel §4 ne é un esempio. Un altro esempio é la teoria dell'entropia dell'azione di $(\mathcal{M}, T)$ rispetto alle scelte di dati iniziali eseguite casualmente rispetto ad una distribuzione di probabilità μ invariante (scelta tale solo per semplificare l'esposizione).

Anche questa nozione prende le mosse dall'opera di Boltzmann e dalla sua ben nota formula esprimente l'entropia termodinamica come proporzionale al numero $\mathcal{N}$ di stati possibili: $S = k \log \mathcal{N}$.

Nella teoria dei sistemi dinamici la nozione di entropia emerge quando si tenta di definire una misura della complessità dell'azione di T su dati iniziali scelti casualmente con distribuzione invariante μ. Immaginiamo di osservare il moto di un punto x scelto a caso, e cioé di costruire la successione $\{T^h x\}$ ottenuta osservando ad ogni istante h in quale insieme $P_j,\ j = 1, 2, \ldots, n$ il punto $T^h x$ viene a trovarsi, se $\mathcal{P} = (P_1, P_2, \ldots, P_n)$ sono n insiemi, privi di punti in comune, nei quali si immagina diviso lo spazio $\mathcal{M}$.

Ad ogni punto $x \in \mathcal{M}$ si potrà associare la "storia" di x rispetto agli insiemi in $\mathcal{P}$: questa è la successione $\underline{\sigma} = (\sigma_h)_{h=0}^{\infty}$ che dice in quale degli

insiemi della collezione $\mathcal{P}$ il punto si viene a trovare ad un dato istante: dunque $\underline{\sigma}$ é definita dalle relazioni:

$$T^j x \in P_{\sigma_j} \qquad j = 0, 1, \ldots \tag{5.1}$$

La complessità di una successione $\underline{\sigma}$ di simboli puó essere misurata dalla quantità di stringhe finite di simboli $\underline{\alpha} = (\alpha_1, \ldots, \alpha_p)$ che "appaiono" in essa.

Precisamente si dice che una stringa di simboli $\underline{\alpha} = (\alpha_1, \ldots, \alpha_p)$ "appare" nella successione $\underline{\sigma}$ se il rapporto fra il numero di volte in cui questa stringa si realizza in $\underline{\sigma}$ con punto di inizio $j < N$ ed il numero N tende ad un limite positivo per $N \to \infty$ (ossia il numero di valori $j < N$, per cui si trova $\sigma_j = \alpha_1, \sigma_{j+1} = \alpha_2, \ldots, \sigma_{j+p-1} = \alpha_p$, diviso per N ha limite positivo per $N \to \infty$). Si dice allora che la stringa $\underline{\alpha}$ é realizzata in $\underline{\sigma}$ con *densità positiva*.

Se ora $\underline{\sigma}$ é la storia di un punto $x \in \mathcal{M}$ si vede subito che la stringa $\underline{\alpha}$ appare con densità positiva in $\underline{\sigma}$ se l'insieme dei valori j per cui $T^j x \in P_{\alpha_1} \cap T^{-1} P_{\alpha_2} \cap \ldots \cap T^{-p+1} P_{\alpha_p} \equiv E$ ha densità positiva, o, in altri termini, se esiste il limite:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \chi_E(T^j x) = \nu(\underline{\alpha}, \underline{\sigma}) > 0 \tag{5.2}$$

che é, in questo caso, un altro modo di scrivere che la frequenza di apparizione di $\underline{\alpha}$ in $\underline{\sigma}$ é positiva, (quì si denota χ_E é la funzione caratteristica di $E = P_{\alpha_1} \cap T^{-1} P_{\alpha_2} \cap \ldots \cap T^{-p+1} P_{\alpha_p}$).

In base ai risultati generali del §2 (teorema ergodico, con $\chi_E = f$) si vede che per μ–quasi tutti i punti x la media (5.2) deve esistere. Se inoltre, come supporremo d'ora in poi per semplificare la discussione, il sistema dinamico agisce ergodicamente su μ, allora il valore del limite in (5.2) non dipende dal punto x scelto, ma solo (con μ–probabilità 1) dalla distribuzione μ di probabilità con la quale viene scelto: e dovrà essere per la definizione di ergodicità (con $\chi_E = f$):

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} \chi_E(T^j x) = \mu(E) = \mu(P_{\alpha_1} \cap T^{-1} P_{\alpha_2} \cap \ldots \cap T^{-p+1} P_{\alpha_p}) \tag{5.3}$$

Dunque le stringhe $\underline{\alpha}$ che appaiono in $\underline{\sigma}$ sono quelle per cui l'insieme $P_{\alpha_1} \cap T^{-1} P_{\alpha_2} \cap \ldots \cap T^{-p+1} P_{\alpha_p}$ ha probabilità positiva.

Ci poniamo ora il problema di contare quante di queste stringhe appaiano nella successione $\underline{\sigma}$, storia di un punto x scelto a caso come detto. E osserviamo che, qualunque sia la definizione che adotteremo, il risultato sarà indipendente da x (con μ probabilità 1) perché per l' ipotesi di ergodicità

e la (5.3) l'insieme delle stringhe che appaiono in $\underline{\sigma}$ non dipende da x. Si suol dire che le stringhe che appaiono in una storia $\underline{\sigma}$ "tipica" per la misura ergodica μ, e le loro frequenze di apparizione, non dipendono dalla storia stessa.

Sia $\mathcal{C}_p$ l'insieme delle stringhe di lunghezza p che appaiono in $\underline{\sigma}$, supposta tipica per l'azione di T su μ. Possiamo dare una misura della quantità di stringhe che appaiono in $\underline{\sigma}$ semplicemente contando il numero $\#(\mathcal{C}_p)$ di elementi di $\mathcal{C}_p$ e calcolando:

$$S_o = \lim_{p\to\infty} \frac{1}{p} \log \#(\mathcal{C}_p) \tag{5.4}$$

che si puó dimostrare esistere sempre (nelle ipotesi considerate); S_o é una misura della ricchezza in stringhe finite della storia $\underline{\sigma}$. Si puó vedere che i moti quasi periodici che abbiamo incontrato nel §4, c.f.r. esempi 2),3), e che possono essere naturalmente considerati come sistemi dinamici agenti su una misura di probabilità naturale μ (che é la misura uniforme sui tori invarianti sui quali tali moti si svolgono) generano storie con $S_o = 0$; e anzi si puó dimostrare di piú e precisamente che la misura μ considerata é in questo caso l'unica misura invariante e così questi moti sono "semplici" in un senso molto forte (e questo é intutivamente soddisfacente). Peró i moti di un punto libero su una superficie a curvatura negativa pure considerati al §4, c.f.r. esempio 3), come anche i moti dei biliardi con pioli rigidi e convessi, c.f.r. esempio 5), danno luogo (come si puó dimostrare) a valori di $S_o > 0$.

Si osservi che la S_o in (5.4) pone sullo stesso piano tutte le stringhe che appaiono in $\underline{\sigma}$ senza tener conto che stringhe di uguale lunghezza possono apparire con frequenza molto diversa.

Perció non ci si contenta di questa misura "ingenua" di complessità e si definisce la complessità in modo diverso che tiene conto di eventuali diversità di frequenza delle stringhe $\underline{\alpha}$ che appaiono nella $\underline{\sigma}$.

Si fissi un numero positivo ε e si pensi di dividere $\mathcal{C}_p$ in due insiemi $\mathcal{C}_p^0$ e $\mathcal{C}_p^1$ in modo che la frequenza totale delle stringhe in $\mathcal{C}_p^0$ sia minore di ε (cioé la somma delle frequenze delle varie stringhe in $\mathcal{C}_p^0$ sia minore di ε).

Sarà possibile fare questo in vari modi e possiamo considerare il "migliore", cioé quello per cui $\mathcal{C}_p^1$ contiene il minimo numero di elementi, pur continuando ad essere vero che $\mathcal{C}_p^0$ ha frequenza totale inferiore ad ε. Questo numero minimo sarà denotato $\mathcal{N}_p(\varepsilon)$, ed é il numero delle stringhe che é veramente necessario tener presenti se si vuol dire che le altre hanno frequenza totale minore di ε.

La "complessità apparente" dell'azione di T su dati scelti casualmente con distribuzione μ e osservati rispetto a $\mathcal{P}$, viene ora definita come:

$$S(\mathcal{P}, T) = \lim_{\varepsilon\to 0} \lim_{p\to\infty} \frac{1}{p} \log \mathcal{N}_p(\varepsilon) \tag{5.5}$$

Si può dimostrare, ma non é immediato, (Shannon-McMillan),che questo limite (5.5) esiste e, poi, che in generale $S(\mathcal{P},T) < S_o$, che aprima vista può sorprendere perchè sarebbe tentante ritenere che le due quantità siano uguali visto che S_o puó anche definirsi come:

$$S_o = \lim_{p\to\infty} \lim_{\varepsilon\to 0} \frac{1}{p} \log \mathcal{N}_p(\varepsilon) \tag{5.6}$$

Questa diversità fra S_o e $S(\mathcal{P},T)$ fà vedere come questa nozione di complessità apparente sia nuova. Ma il vero motivo per cui la nozione é interessante sta nel fatto che, contrariamente a quanto si potrebbe forse credere ingenuamente, il numero $S(\mathcal{P})$ *non cresce* indefinitamente al crescere del numero di elementi in $\mathcal{P}$, ossia eseguendo osservazioni piú dettagliate: aumentando il dettaglio con cui si studia il moto la complessità apparente spesso non puó superare un certo limite finito.

Si puó, invece, dimostrare che, nei sistemi dinamici regolari che qui consideriamo:

$$S(T) = \sup_{\mathcal{P}} S(\mathcal{P},T) < \infty \tag{5.7}$$

Inoltre se T é supposta invertibile e se $\mathcal{P}_o$ é tale che la diversità delle storie di x e x' per T o T^{-1} implichi $x \neq x'$, se cioé $\mathcal{P}_o$ é abbastanza fine e T abbastanza non banale perché a storie diverse corrispondano punti diversi che le generano, allora (Kolmogorov, Sinai):

$$S(T) = S(\mathcal{P}_o, T) \tag{5.8}$$

La quantità definita in (5.7), massimo valore delle complessità apparenti, viene chiamata "entropia" o "complessità" dei moti di T con dati iniziali casuali rispetto a scelte di x con distribuzione μ.

Ovviamente $S(\mathcal{P},T) \leq S_o$ per ogni $\mathcal{P}$ e quindi i moti quasi periodici prima considerati e intesi come sistemi dinamici che agiscono su una misura invariante hanno entropia nulla.

Peró i sistemi associati al moto libero su una superficie a curvatura negativa e i biliardi con pioli convessi e rigidi, considerati come sistemi dinamici agenti sulle rispettive misure μ di Liouville, non solo sono tali che $S(\mathcal{P}) > 0$ ma anche sono tali che $S(\mathcal{P},T) > 0$ per ogni $\mathcal{P}$ che sia non banale (cioé tale che uno almeno degli insiemi in $\mathcal{P}$ abbia probabilità $\neq 0,1$). In questi sistemi i moti non possono apparire semplici e hanno sempre complessità apparente positiva (a meno che $\mathcal{P}$ sia banale). Sono moti così complicati che anche rinunciando ad osservare alcuni dettagli, la complessità puó si diminuire ma non diviene mai nulla: i sistemi dinamici che possiedono questa proprietà sono chiamati "K-sistemi".

La nozione di entropia si estende senza difficoltà ai sistemi dinamici a piú dimensioni incontrati nel §4, c.f.r. esempio 5), e un' analoga proprietà é valida per i sistemi dinamici che si incontrano nella meccanica statistica, cui si é pure accennato al §4. Gran parte delle misure di probabilità invarianti per traslazione che descrivono fasi pure risultano descritte da distribuzioni di probabilità μ sulle quali le traslazioni $\underline{T}$ agiscono in modo che le configurazioni x del sistema, osservate all'interno di un reticolo di scatole che riempiono lo spazio, generano successioni a complessità positiva. Così, comunque poco in dettaglio si osservi la distribuzione delle particele in un gas, questa apparirà a complessità > 0, forse piccola ma *mai* nulla. In questi sistemi si puó anche vedere che esiste una stretta relazione ed una sostanziale identità fra la nozione di entropia ora introdotta (per l'azione delle traslazioni spaziali) e l' entropia della termodinamica, che é legata alla formula di Boltzmann. Questo giustifica il nome dato alla (5.8).

La nozione di entropia é importante anche nella teoria degli "isomorfismi" fra sistemi dinamici, cioé per riconoscere se due sistemi dinamici possano considerarsi come ottemuti l'uno dall'altro a mezzo di un cambiamento di coordinate almeno per quel che concerne la loro azione su date distribuzioni invarianti di probabilità (ossia per riconoscere se siano collegati da un cambiamento di coordinate con "singolarità" eventualmente presenti ma concentrate su insiemi di misura nulla per le distribuzioni di probabilità considerate). In questo modo é stato possibile, grazie in particolare alle idee innovative del lavoro di Ornstein, stabilire che sistemi dall'apparenza assai diversa erano riducibili gli uni agli altri, almeno per quel che concerneva la loro azione su certe misure invarianti di probabilità.

E' opportuno arrestare quì, dopo la rapida analisi delle idee base della nozione di entropia, questa breve introduzione alla teoria ergodica rimandando il lettore interessato alla letteratura specializzata avvertendo che, come energe da quanto detto sopra, si tratta di una letteratura in pieno sviluppo e quindi non risulta ancora ben organizzata dal punto di vista espositivo, e sia spesso dedicata a problemi di natura estremamente tecnica.

Bibliografia

a) fondamenti

[1]L. Boltzmann: in "Wissenshaftlichen Abhandlungen", Chelsea, New York, 1968, "Über die eigenshaften monozyklischer und anderer damit verwandter systeme", p.122–152, originale in "Crelles Journal, 98,68–94,1884.
[2]L. Boltzmann: in "Wissenshaftlichen Abhandlungen", Chelsea, New

York , 1968, "Entgegnung auf die warme theoretischen betrachtungen des Hrn. Zermelo", p.567–578, originale in Wied. Ann., 57, 773–784, 1896.
[3]L.Boltzmann: "Vorlesungen uber gas theorie", I, II, Leipzig, Barth, 1895–1898 (trad. inglese: "Lectures on gas theory", ed. S. Brush, Berkeley Univ. Press, 1964).
[4]P. e T. Ehrenfest: "The conceptual foundations of the statistical approach in Mechanics, Cornell Univ. Press, 1959 (trad. inglese dell'originale "Begriffliche grundlagen der statistische auffassung in der mechanik", IV, Encyklopädiae. der Mathematischen Wiss., Teubner, Leipzig, 1911).
[5]S.G. Brush: "Kinetic theory", Pergamon Press, New York, 1967–72.
[6]S.G. Brush: "The kind of motion we call heat", I, II, North Holland, Amsterdam, 1976.
[7]N. Krylov: "Works on the foundations of statistical mechanics", Princeton Univ. Press, Princeton, 1979.

b) Strumenti matematici di base

[8]M. Kac: "Stocastica", in Enciclopedia del 900, Enciclopedia Italiana, Roma, 19**.
[9]G.C. Rota, P.S. Kung, "Probabilità", in Enciclopedia del 900, Enciclopedia Italiana, Roma, 19**.
[10]P. Halmos: Lectures on ergodic theory, Publications of the mathematical society of Japan, Tokio, 1956.
[11] P. Billingsley: Ergodic theory and information", Wiley, New York, 1965
[12]P. Walters: "Ergodic theory", Lecture notes in Mathematics, vol. 458, Springer Verlag, Berlin, 1975.
[13]A. Kintchin: "Mathematical foundations of information theory", Dover, New York, 1957.
[14]V.I. Arnold, A. Avez,"Ergodic problems of classical mechanics", Benjamin, New York, 1965
[15]G. Gallavotti: "Aspetti della teoria ergodica, qualitativa e statistica del moto", Quaderni dell' Unione Matematica Italiana, vol. 21, Pitagora, Bologna, 1981.

Caos.

Voce per il Dizionario delle Scienze Fisiche
dell' Enciclopedia Italiana,
Giovanni Gallavotti
Agosto 1989.

§1 I moti regolari.

La concezione antica del moto come sempre scomponibile in moti circolari uniformi, nata dalle osservazioni astronomiche, raggiunse il suo massimo sviluppo nell' *Almagesto* di Tolomeo. In termini moderni un moto è scomponibile in n moti circolari uniformi se ogni coordinata che lo descrive, o più in generale, ogni grandezza osservabile A, osservata al variare del tempo t può essere scritta come:

$$A(t) = f_A(\omega_1 t, \dots, \omega_n t) \tag{1}$$

ove $f_A(\varphi_1, \dots, \varphi_n)$ è una funzione periodica con periodo 2π di n angoli $\varphi_1, \dots, \varphi_n$; le quantità $T_1 = 2\pi\omega_1^{-1}, \dots, T_n = 2\pi\omega_n^{-1}$ sono gli n *periodi* del moto e i loro inversi $\nu_1 = 1/T_1, \dots, \nu_n = 1/T_n$ sono le *frequenze*.

Si possono quindi immaginare n cerchi di raggio 1 e che lo stato del sistema sia determinato all' istante t dalla posizione, su questi cerchi, di n punti che ruotano con velocità angolari rispettive $\omega_1, \dots, \omega_n$. Questi moti sono ora chiamati *quasi periodici*, a n periodi o a n frequenze.

Se i cerchi vengono considerati di raggio R si scriverà v_i/R in luogo di ω_i, con v_i= velocità dell' i-mo punto, e si potrà considerare anche il moto rettilineo uniforme come un caso limite di moto circolare uniforme.

L' *Almagesto* di Tolomeo è dedicato alla determinazione delle funzioni periodiche f per le coordinate geocentriche degli astri, a partire dai dati sperimentali sui moti celesti: è noto che l' intero sistema solare, più le stelle fisse, potè essere rappresentato da Tolomeo a mezzo di 52 moti circolari

uniformi (classificati in deferenti ed epicicli) con al più 5 per un singolo pianeta (Marte).

Poichè una funzione multiperiodica f_A è determinata univocamente dal suo sviluppo di Fourier, si può dire che l' astronomia greca consistè nella determinazione empirica, cioè dedotte dalle osservazioni sperimentali, dei coefficienti di Fourier delle funzioni periodiche che, via la (1), descrivono il moto degli astri.

La fisica di Newton introdusse un nuovo fondamentale elemento: la legge di gravitazione universale, che permise di determinare *a priori*, sulla base dai pochi dati iniziali, le funzioni f_A, e di limitare, *a priori*, il numero n delle frequenze *indipendenti*, (ossia tali che $\sum k_j \nu_j = 0$, con k_j interi, implichi $k_j \equiv 0$), con il numero di gradi di libertà. Concettualmente, dunque, il moto restava come nella concezione greca, ma ora si possedeva un algoritmo per calcolare le funzioni f_A, cioè si aveva una teoria *a priori* del "Sistema del Mondo": così Laplace dimostrò con il suo *Trattato di Meccanica Celeste*.

La fisica newtoniana forniva inoltre la possibilità di una teoria dei moti dei sistemi meccanici di natura non astronomica. E la teoria dei moti dei sistemi più semplici, come oscillatori unidimensionali (pendolo), sistemi di oscillatori armonici (corda vibrante), giroscopi, moti senza attrito su ellissoidi, risultarono ancora concepibili come moti risultanti di moti circolari uniformi: nella meccanica contemporanea i sistemi di questo tipo si chiamano *integrabili (per quadrature)*, v. teoria delle perturbazioni.

§2 I moti caotici.

E' solo con Boltzmann e Poincaré, alla fine del '800 che appare chiaro che i moti circolari uniformi non possono esaurire tutte le possibilità e cercare, sempre, di ridursi ad essi sarebbe pericolosamente riduttivo.

E tuttavia, fino ad oltre la metà del '900 la tendenza a cercare di ridurre i problemi a problemi di moti quasi periodici (con la conseguente ricerca dei *modo normale*, v., e delle *armonica fondamentale*) è stata assai forte.

Un esempio paradigmatico è offerto dalla teoria di Landau della turbolenza nei fluidi. La teoria propone che, al crescere di un parametro di controllo che misura la intensità della forza che tiene in moto stazionario il fluido viscoso (numero di *Reynolds*), avvenga che il moto del fluido passi da un moto indipendente dal tempo in ogni punto, ad un moto periodico, poi ad moto quasi periodico a due periodi, poi a tre, quattro, *etc.* Quando il numero delle frequenze presenti sarà sufficientemente grande, allora il fluido apparirà dotato di un moto stazionario (v.), ma irregolare, ossia turbolento.

Ma i tentativi di determinare quantitativamente i periodi in questione

hanno incontrato grandi difficoltà e hanno, in ultima analisi, condotto ad un ripensamento profondo di tutto il modo di affrontare il problema della turbolenza sulla base della concezione quasi periodica dei moti. Ed è emersa in modo inequivocabile la necessità dell' uso di moti non quasi periodici per la rappresentazione corretta, non solo dei fenomeni di turbolenza, ma anche di fenomeni di natura completamente diversa (quali i moti degli assi di satelliti, o i moti di sistemi meccanici semplici, o i modelli di evoluzione di popolazioni o di reazioni chimiche).

I moti non quasi periodici verranno qui chiamati *moti caotici* anche se a volte si preferisce riservare questo nome a moti che abbiano specifiche proprietà addizionali.

Volendo dunque dare una definizione precisa si dirà che un moto di un sistema confinato in un volume finito (per semplicità ci limitiamo a questi moti) è *caotico* se verifica un' equazione differenziale ordinaria o alle derivate parziali e, quindi, *deterministica* (perchè supporremo sempre che le equazioni che descrivono i sistemi ne determinino univocamente i moti, a partire dai dati iniziali), e se esiste almeno una grandezza osservabile A (cioè una funzione delle coordinate che descrivono lo stato del sistema) che osservata al variare del tempo è una funzione $t \to A(t)$ che *non* è funzione quasi periodica del tempo, o almeno ha un numero di frequenze indipendenti superiore al numero di gradi di libertà. L' uso dell' aggettivo "caotico" è motivato dall' insieme di proprietà che di solito, nelle applicazioni teoriche e sperimentali, sono concomitanti a questa proprietà di base.

Prima di procedere ad un esame più dettagliato delle proprietà dei moti caotici è bene chiarire due obiezioni che spesso vengono sollevate a questo punto.

Il lettore potrebbe obiettare che la differenza fra moti quasi periodici e moti caotici è, in pratica assai evanescente: la teoria degli sviluppi di Fourier fa vedere che ogni funzione può essere espressa via un integrale di Fourier. E le funzioni quasi periodiche sono una classe particolare in cui l' integrale di Fourier si riduce ad una somma estesa ad un insieme di armoniche (formante un insieme denso di punti sulla retta) della forma $(k_1\nu_1 + \ldots + k_n\nu_n)$ con k_j interi, e con la proprietà aggiuntiva che la massima parte delle armoniche ha di solito piccola ampiezza e quindi è, praticamente, non osservabile. Ma è chiaro che, siccome gli integrali sono approssimabili con arbitraria precisione da somme, è sempre possibile rappresentare una funzione $A(t)$ come quasi periodica con una appprossimazione prefissata. In generale, però, per fare questo occorre un numero N di frequenze fondamentali indipendenti tanto più grande quanto maggiore è la precisione desiderata; ovvero occorre che moltissime armoniche abbano ampiezza non trascurabile, tante più quanto maggiore è la precisione desiderata. Nella visione del moti come moti quasi periodici il numero di frequenze è però limitato dal numero di

gradi di libertà: e quindi nei sistemi a pochi gradi di libertà la distinzione fra moti quasi periodici e no è chiara. Nei sistemi a infiniti gradi di libertà, quale un fluido, la distinzione è meno netta (pur essendo ancora tale da un punto di vista strettamente matematico): a meno che il sistema sia un sistema dissipativo. d In questi casi ci si aspetta (come nella teoria di Landau) che il numero di frequenze indipendenti dei moti quasi periodici, anche in regime di turbolenza assai sviluppata, sia molto più piccolo del numero di gradi di libertà, e anzi finito anche se grande e crescente gradualmente al crescere delle forze applicate. E' infatti ben noto che assai di rado si considerano nelle applicazioni modelli di fluido realmente descritti da un continuo e si preferisce studiare modelli a un numero finito di gradi di libertà, perchè più facilmente trattabili a mezzo di elaboratori elettronici: di solito a un fissato valore delle forze esterne si rivela inutile aumentare il numero di gradi di libertà del modello oltre un certo limite. E dunque si possono immaginare esperimenti che mettano in evidenza eventuali quasi periodicità.

D' altra parte l' esistenza di moti non periodici si può manifestare in un modo un pò paradossale (che fu il fulcro della polemica fra Boltzmann e i suoi critici, e quindi solo *a posteriori* chiaro). In Boltzmann i moti non periodici appaiono attraverso l' ipotesi *ergodica*, v., in una forma a prima vista contradditoria. Quest' ipotesi, secondo Boltzmann, si formula immaginando la spazio delle fasi, di un sistema di N particelle confinate in un volume V, come suddiviso in cellette, v., ognuna delle quali descrive uno stato microscopico del sistema. La evoluzione temporale (osservata a intervalli di tempo microscopici, ma non troppo piccoli, v. MSC), trasforma successivamente una celletta in un' altra senza mai che una celletta torni alla celletta iniziale se non dopo essere stata trasformata in *tutte* le altre di uguale energia.

Dunque, nei casi in cui tale ipotesi è ragionevole (e Boltzmann pensava certamente ai gas rarefatti con interazioni a cuore repulsivo) il moto di ogni dato iniziale (celletta) sarebbe addirittura periodico!. Però Boltzmann stimò il periodo come più grande, e di moltissimi ordini di grandezza, dell' età dell' universo, per un gas di poche decine di particelle. Perciò una tale periodicità è priva di qualsiasi interesse e irrilevante. Il vero problema è dunque se su scale di tempo molto più piccole dell' età dell' universo, dell' ordine del giorno o del millesimo di secondo, il moto possa apparire o no, nelle approssimazioni che ci interessa di raggiungere, un moto semplice, periodico o quasi periodico: la risposta di Boltzmann, implicita nelle proprietà dell' equazione di Boltzmann e nella lunga polemica con i suoi critici, fu negativa.

Poincaré vide chiaramente la necessità di moti non quasi periodici per una corretta descrizione delle proprietà, anche se solo qualitative, dei moti dei

sistemi anche di gran lunga più semplici dei gas, come ad esempio i moti del problema dei tre corpi, v., della Meccanica Celeste. Ma la piena comprensione di questi fatti restò patrimonio di pochi fino a che lo sviluppo degli elaboratori elettronici rese il fenomeno facilmente osservabile e comprensibile, anche da parte di chi avesse scarso interesse per la matematica.

§3 *Caos nelle previsioni meteorologiche.*

È ora chiaro, ed accettato, che i moti quasi periodici sono, soprattutto nei sistemi dissipativi ma (nel senso ingenuo della integrabilità per quadrature) anche nei sistemi conservativi, un' eccezione, anche se la loro semplicità fa sì che lo studio offra importanti spunti alla comprensione dei moti caotici, che spesso si possono immaginare come generati da moti quasi periodici instabili.

Per brevità di esposizione concentreremo la discussione sui sistemi dissipativi a un numero finito (ma arbitrariamente grande) di gradi di libertà, accennando solo marginalmente a proprietà analoghe dei sistemi conservativi.

Il primo sistema che è stato descritto come animato da moti caotici, nell' ottica di una critica ai moti quasi periodici e alla sua rilevanza in questioni di grande importanza applicativa (quelle delle previsioni meteorologiche), è stato il *sistema di E. Lorenz*, (1963), che nel descriverlo ne mise esplicitamente in luce gli aspetti caotici e ne rilevò l' esistenza non come una curiosità del caso particolare da lui studiato, bensì come una proprietà assai comune nei sistemi deterministici, introducendo al tempo stesso alcune tecniche molto utili per descrivere *ordinatamente* il moto caotico.

Il modello di Lorenz è un sistema descritto da equazioni dedotte dalle equazioni della convezione(v.), con opportune condizioni al bordo e con approssimazioni (assai grossolane). L' intuizione di Lorenz fu che, sebbene queste equazioni non potessero essere considerate una buona approssimazione dal punto di vista quantitativo, tuttavia fossero utilizzabili per comprendere alcune proprietà qualitative dei sistemi cui si è realmente interessati: e ne risultò l' idea dell' ubiquità dei moti caotici. Il sistema di equazioni è estremamente semplice; è a tre incognite x, y, z e si scrive:

$$\begin{aligned} \dot{x} &= -\sigma(x-y) \\ \dot{y} &= -xz + rx - y \\ \dot{z} &= xy - bz \end{aligned} \tag{2}$$

con $\sigma = 10.$, $b = 8/3$ e r parametro variabile *numero di Prandtl*), che misura l' intensità delle forze che tengono in moto il sistema: infatti se $r = 0$ si

vede facilmente che la quantità $x^2/\sigma + y^2 + z^2$ tende a zero per $t \to \infty$, cioè il sistema tende banalmente allo stato di equilibrio $x = y = z = 0$. Se $r \neq 0$ allora si può vedere che il sistema evolve in modo che il punto (x, y, z) entra in una sfera con centro $x_0 = 0, y_0 = 0, z_0 = r - 1$ e raggio abbastanza grande in un tempo finito e vi resta indefinitamente, qualsiasi sia il dato iniziale del moto.

Dunque i moti del sistema sono confinati e si vuole studiarne il comportamento asintotico per $t \to \infty$. Si trova che, per r piccolo, ogni dato iniziale evolve semplicemente verso lo stato di equilibrio $x = y = z = 0$. Ma, al crescere di r (per $r > 1$), si osserva che questo stato di equilibrio non descrive più il comportamento asintotico di uno stato iniziale scelto a caso: solo specialissimi dati iniziali evolvono tendendo a $x = y = z = 0$ (fra questi quelli per cui $x = y = 0$ e z è arbitrario). Il dato iniziale generico, invece, evolve verso *uno* degli altri due punti di equilibrio $x = y = \pi, \sqrt{b(r-1)}$, $z = r - 1$, almeno finchè r è abbastanza piccolo. Continuando a far crescere r si raggiunge un valore critico r_c e per, $r > r_c$, un dato iniziale preso a caso (in una sfera e con distribuzione uniforme, ad esempio) non evolve nè verso l' origine, nè verso gli altri due punti fissi (che dunque rappresentano ora stati di equilibrio instabile).

Pertanto, per r grande, uno stato iniziale evolve senza fermarsi mai: però il suo moto non è nè periodico, nè quasi periodico ma appare (Lorenz) molto irregolare, pur svolgendosi asintoticamente su un sottoinsieme dello spazio dei dati che ha misura nulla. Lorenz riuscì a porre "ordine" nella descrizione di questi moti. Osservò che, dopo un transiente iniziale, il valore di una coordinata (la z, nella sua scelta) misurata a istanti $t_1, t_2, \ldots$ scelti in modo che la $z(t)$ avesse agli istanti di osservazione un massimo relativo, dava luogo a successioni $z_1, z_2, \ldots$ assai diverse per diversi dati iniziali ma legate fra loro e con ottima approssimazione, (anche se non esattamente), da una legge della forma $z_{n+1} = f(z_n)$, ove f è una funzione dedotta dall' esame dei dati. I valori delle z variano fra un minimo z_{min} ed un massimo z_{max} in modo che ponendo $x = (z - z_{min}/(z_{max} - z_{min})$ la f può essere considerata come una trasformazione dell' intervallo $[0, 1]$ in se stesso. In *prima approssimazione* la f può essere scritta analiticamente in modo semplice (e le proprietà che se ne deducono non dipendono dalla semplicità della funzione ma solo dal fatto che ha ovunque modulo della derivata > 1). La rappresentazione è:

$$x_{n+1} = 2x_n \mod 1 \tag{3}$$

che ha un grafico semplice *a tenda*, con una derivata a modulo costante $(= 2)$ discontinua in $x = 1/2$ ove cambia segno.

La (3) ha le seguenti due proprietà:

1) dipendenza sensibile dai dati iniziali: la differenza $x'_n - x_n$ fra due succes-

sioni con dati iniziali x'_0, x_0, anche molto vicini, cresce come 2^n, cioè esponenzialmente con il tempo (poichè il tempo che intercorre fra due massimi successivi ha una durata media, empiricamente misurata, che non dipende dai punti iniziali). Naturalmente questo resta vero solo finchè $|x'_n - x_n|$ è piccolo, in modo che x'_n e x_n cadano nella stessa metà dell' intervallo $[0, 1]$; poi le due successioni evolvono in modo diverso e $|x'_n - x_n|$ mantiene un valore medio dell' ordine di 1 (cioè dell' ordine della massima differenza possibile).
2) complessià (o entropia, v.) positiva: se si sceglie un dato iniziale a caso $x \in [0, 1]$, con distibuzione uniforme, e se si costruisce la *storia* $\underline{\sigma}(x_0)$ di x_0 sugli insiemi $P_1 = [0, 1/2)$ e $P_2 = [1/2, 1]$, cioè se si definisce la successione $\underline{\sigma}(x_0) = (\sigma_0, \sigma_1, \ldots)$ con $\sigma_i = 1, 2$ tali che $f^i(x_0) \in P_{\sigma_i}$, allora la probabilità che il dato iniziale x_0 sia tale che i primi k elementi della successione $\underline{\sigma}(x_0)$ coincidano con una stringa data, arbitrariamente, è 2^{-k}.

In altre parole il moto di un x_0 assegnato a caso in $[0, 1]$, con distribuzione uniforme, produce successioni di simboli $\sigma_0, \sigma_1, \ldots$ che hanno la stessa statistica di quelle prodotte dal lancio di un dado perfetto, a due facce (v. successioni a frequenza definita). Tali successioni sono a entropia (v.), o complessità, positiva. Inoltre se il dato iniziale delle equazioni viene scelto con distribuzione uniforme in una sfera prefissata, si trova (sempre empiricamente) che se si attende abbaztanza prima di iniziare le osservarezioni (cioè dopo un transiente abbastanza lungo), la prima delle coordinate x, cioè la x_0, si trova nell' intervallo $[0, 1]$ con una distribuzione di probabilità che, pur non essendo uniforme, ha tuttavia una distribuzione di probabilità descritta da una densità in $[0, 1]$. Quindi, anche scegliendo i dati iniziali a caso in una sferetta, le successioni x_i che si osserveranno avranno la stessa statistica dei lanci di un dado perfetto.

E' chiaro che le due proprietà appena descritte suggeriscono di classificare il moto in esame come un *moto caotico*. La 1) ci dice che il moto è, da un punto di vista pratico, imprevedibile a dispetto del suo determinismo: a causa della divergenza esponenziale delle traiettorie, diviene necessaria, per la predizione del valore x_n in termini di x_0 e per n grande, una conoscenza del dato iniziale x_0 con una precisione in pratica irrealizzabile (già per valori di n di qualche decina se $|f'| = 2$, come nel caso in esame; ma comunque, anche se 2 fosse sostituito con un numero vicino a 1, seppur di poco, non si avrebbe un cambiamento qualitativo dell' ostacolo alla predicibilità).

La 2) ci dice che il moto è talmente caotico da poter essere utilizzato, almeno in linea di principio, per produrre successioni di numeri a caso (come quelli che si ottengono lanciando un dado perfetto).

E' possibile mostrare che un moto quasi periodico non può generare moti che manifestino le due proprietà appena descritte: dunque il sistema di Lorenz, per r grande (si trova per $r \geq 82$), è tale che i moti da esso descritti

sono caotici, e in senso assai chiaro, (salvo dati iniziali eccezionali, formanti un insieme di misura nulla nello spazio dei dati iniziali, fra cui i dati con $x = y = 0$).

§4 Teoria di Ruelle Takens: gli attrattori strani.

Pochi anni dopo il lavoro di Lorenz, che non destò particolare risonanza, apparve il lavoro, indipendente, di D. Ruelle e F. Takens: prendendo le mosse da una critica alla concezione dei moti dei fluidi turbolenti come moti quasi periodici a molte frequenze pervennero ad un' affermazione che sorprese molti. La sua essenza è "anche supponendo che un modello per il moto di un fluido (cioè un' equazione differenziale dissipativa di evoluzione) dia luogo a moti che sono asintoticamente quasi periodici con almeno *tre* frequenze indipendenti, dovrebbe accadere che modificando, anche di pochissimo, le equazioni le nuove equazioni diano luogo a moti asintoticamente non più quasi periodici, bensì caotici perchè descritti da (quello che fu chiamato) un *attrattore strano*". In questa analisi si insiste sul fatto che si considerano sistemi dissipativi perchè i sistemi conservativi, descritti da equazioni hamiltoniane, sono dotati di particolari proprietà, per cui i risultati di genericità del lavoro di Ruelle Takens non si applicano, almeno non banalmente, ad essi.

Questo significa che il modello modificato mostrerebbe moti in cui traiettorie corrispondenti a dati iniziali vicini si separano con legge esponenziale (come nel modello di Lorenz) e che osservati ad intervalli di tempo scelti con opportuni criteri (ad esempio ogni volta che una coordinata assume un dato valore, o un valore di massimo relativo) possono dare luogo a successioni di risultati a frequenze definite (v. Frequenza) e a statistica casuale (come nel caso estremo del modello di Lorenz) codificabili, a mezzo di un opportuno *codice*, (cioè cambiamento di coordinate), in successioni di numeri casuali e indipendenti (v. Entropia e Informazione, Bernoulli).

Matematicamente un *attrattore* per le soluzioni $x \to S_t x$ di un' equazione differenziale dissipativa (il lavoro di Ruelle Takens si applica infatti a tutti i sistemi dissipativi, anche se fu motivato dalla teoria della turbolenza) è un insieme chiuso A dotato di un intorno U aperto tale che ogni dato iniziale $x \in U$ evolve in modo che la sua distanza da A tenda a 0 per $t \to \infty$; si richiede anche che A sia minimale fra gli insiemi che godono di questa proprietà (per evitare casi banali). Se l' avvicinamento di $S_t x$ ad A è abbastanza rapido, si può trovare un punto $y \in A$ tale che $d(S_t x, S_t y) \to 0$: allora y è la *proiezione* di x su A e ogni moto che origina in U si può pensare approssimato da un moto in A ad esso asintotico.

Più in generale, ci si pone il problema di come descrivere il moto di un punto $x \in U$, cioè nel *bacino d' attrazione di* A, in termini di proprietà

di A e dei moti su A. Se l' attrattore ha opportune proprietà può essere possibile associare a tutti i punti $x \in U$, salvo eventualmente un insieme di misura nulla, una misura di probabilità μ definita *su* A che descrive la *statistica dei moti* dei moti originanti in U. Nel senso che il valore medio di un' osservabile qualunque G, su un moto $x \to S_t x$ con dato iniziale in $x \in U$, è dato da:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T G(S_t x)dt = \int_A G(x)\mu(dx) \tag{4}$$

Se questo avviene e se, *inoltre*, il sistema dinamico (A, μ, S_t) è, per $t > 0$, isomorfo ad uno schema di Bernoulli, o almeno è ergodico ed ha entropia positiva, (v. Entropia e Informazione, Sistemi Dinamici), si dice che l' attrattore A ha la proprietà di *stranezza*. Questo secondo requisito esclude dalla qualifica di *strani* gli attrattori che siano costituiti da orbite periodiche o tori su cui ha luogo un moto quasi periodico.

La definizione precisa di *attrattore strano* varia nella letteratura e spesso se ne adottano definizioni matematicamente convenienti ma troppo restrittive per le applicazioni. La definizione qui adottata, cioè la validità della (4) e della successiva proprietà del sistema dinamico (A, μ, S_t), è quindi solo indicativa.

Il fatto che perturbando anche di pochissimo un modello che produce moti quasi periodici a più di due frequenze indipendenti si possa ottenere (eRuelle Takens nel loro lavoro dimostrano anche che, in un senso opportuno, "di solito" si ottenga) un sistema con attrattori strani costringe a ripensare l' idea della turbolenza come descritta da moti quasi periodici a molte frequenze. In assenza di un principio fisico che imponga che i moti di un fluido, o più in generale, di un sistema descritto da un' equazione differenziale (quindi "deterministico") siano quasi periodici non ci si può aspettare che i moti quasi periodici (a più di due frequenze) siano la regola. Infatti i modelli che descrivono fenomeni fisici devono essere insensibili a piccole variazioni dei modelli stessi, purchè siano compatibili con i principi fondamentali (quali le leggi di conservazione macroscopiche e microscopiche); e poichè nessun principio impone che i modelli diano luogo a moti quasi periodici, non è naturale fondare su di essi teorie; almeno non nei casi in cui si può mostrare che piccole modifiche del modello conducono a modelli con moti caotici (come nel caso di modelli di fenomeni dissipativi con moti quasi periodici a più di due frequenze).

§5 *Conseguenze sperimentali del nuovo punto di vista.*

Il caos dovrebbe dunque essere la *regola* in tutti i fenomeni abbastanza complessi da richiedere, per essere descritti a mezzo di moti quasi periodici, più di due frequenze indipendenti.

Si vede il motivo dell‘ impatto del lavoro di Ruelle Takens sulla ricerca: fino al 1970 i lavori sperimentali sui fluidi, sulle reazioni chimiche e, in generale, sui problemi di evoluzione in sistemi dissipativi non avevano prestato grande attenzione ai moti poco disordinati. Le osservazioni cercavano direttamente di studiare e introdurre un pò di ordine nei moti molto disordinati, di solito cercando di interpretarli come moti quasi periodici. Il lavoro di Ruelle Takens suggeriva un gran numero di esperimenti volti a controllare la impossibilità, nella maggior parte dei sistemi, dei moti quasi periodici a più di due frequenze. Si tratta di esperimenti di gran lunga più facili di quelli da eseguire per lo studio dei sistemi con moti molto disordinati; quindi si tratta di esperimenti che possono essere svolti con grande accuratezza e riproducibilità.

Gli esperimenti possono vertere su sistemi dissipativi reali, fluidi, reazioni chimiche *etc*, ovvero possono essere esperimenti numerici su elaboratori elettronici. Consistono nel far variare, in modo semplice in funzione di un parametro di controllo, la forza che agisce su un sistema per mantenerlo in moto e osservare il comportamento asintotico dei moti che conseguono, scegliendo il dato iniziale a caso in una prefissata regione dello spazio delle fasi. Partendo da valori iniziali del parametro di controllo, cui corrispondono moti asintotici semplici, come stati di equilibrio o moti periodici, si fa variare il parametro in modo che i moti siano via via più complessi. Così si può passare ad esempio da un moto che tende ad una posizione di equilibrio ad un moto che tende ad un moto periodico, ad uno che tende ad un moto quasi periodico a due frequenze e poi finalmente si può osservare se effettivamente appaiano moti a tre frequenze o più, ovvero se il moto cambi carattere divenendo caotico.

E‘ storia nota che nel breve volgere di qualche anno l‘ evidenza per la correttezza dello schema di Ruelle Takens è divenuta schiacciante: sono assai rari i casi in cui è possibile osservare tre frequenze indipendenti (anche se pare che qualcuno sia stato osservato): ciò non significa che l‘ evoluzione della complicazione di un moto asintotico segua sempre e necessariamente lo schema di una poi due frequenze e poi caos. La critica di Ruelle e Takens all‘ evoluzione della complicazione di un moto asintotico, attraverso il semplice e graduale aumento delle frequenze indipendenti, non ha mai sostenuto che questa fosse l‘ unica possibilità (ne illustreremo altre fra poco). Ha invece sostenuto che nei casi in cui il moto asintotico evolvesse, al variare del parametro di controllo, da uno stadio quasi periodico a due frequenze ad uno più complicato non ci si dovrebbe attendere che il moto più complicato sia semplicemente un moto a tre o più frequenze ma piuttosto un moto caotico.

§6 Il principio variazionale di Ruelle.

La situazione nei sistemi conservativi, cioè descritti da equazioni di tipo hamiltoniano è diversa: questi sistemi sono intrinsecamente dotati di simmetrie speciali (legate alla canonicità delle equazioni del moto): ad essi la teoria di Ruelle e Takens non si può applicare direttamente perchè una delle ipotesi essenziali è la possibilità di considerare come buon modello per un dato fenomeno una *qualunque* piccola modifica di un buon modello prefissato. Se si considera un sistema conservativo, invece, non si potranno permettere, ovviamente, modifiche alle equazioni che ne alterino il carattere hamiltoniano rendendole non più conservative.

Tuttavia la nozione di moto caotico ha senso indipendentemente dalla natura non hamiltoniana delle equazioni che lo descrivono e, quindi, moti caotici sono possibili (e in un certo senso addirittura più comuni) in sistemi conservativi: anzi da tempo erano noti semplici e importanti esempi (come il moto geodetico nelle geometrie non euclidee a curvatura negativa costante (Hopf, Morse): che però erano restati nell‘ ambito della matematica).

In prima approssimazione si può dire che non c‘ è grande differenza fra sistemi conservativi e sistemi dissipativi. Nei primi non si può parlare di "attrattori"; se esistono sono piuttosto banali e coincidono con tutto lo spazio delle fasi di data energia (nei sistemi ergodici). Nei secondi si deve immaginare lo spazio delle fasi come una unione di insiemi invarianti A sui quali sono definite distribuzioni di probabilità μ che permettono di descrivere, via la relazione (4), la statistica dei moti con dato iniziale x scelto a caso con una naturale distribuzione di probabilità su A qualed, ad esempio, la misura indotta (*misura condizionale*) su A dalla misura di Liouville. Dunque questi insiemi invarianti giocano lo stesso ruolo degli attrattori dei sistemi dissipativi: però viene a mancare la nozione di bacino di attrazione e quindi all‘ interno di una regione aperta anche piccola è possibile che esistano insiemi di dati iniziali, con misura positiva, evolventi con moto caotico perchè si trovano su un insieme A su cui il moto è caotico e insiemi di misura pure positiva di dati evolventi con moto regolare, quasi periodico; inoltre μ può dipendere in modo sensibile dad$A$.

Però il moto che ha luogo sugli attrattori A dei sistemi dissipativi non va pensato come diverso da quello che ha luogo sugli insiemi invarianti A minimali dei sistemi conservativi: ad esempio la misura μ che descrive le proprietà statistiche dei moti con dati a caso rispetto alla misura di volume (via la (4)) gioca nei sistemi dissipativi il ruolo della misura indotta su A dalla misura di Liouville nei sistemi conservativi.

Questa analogia è assai profonda: come mostrato da Ruelle in vari casi, semplici ma istruttivi, di sistemi deterministici dotati di attrattori strani è possibile verificare che la misura μ che descrive la statistica dei moti $t \to$

$S_t x$ dei dati iniziali scelti a caso nel bacino di attrazione con distribuzione di probabilità proporzionale alla misura di volume, verifica una proprietà variazionale (v. principio variazionale, Limite Termodinamico) analoga a quella verificata dalla misura di Liouville, sulla superfice di energia costante, nella MSC.

Precisamente, fissato $t > 0$ arbitrariamente, fra tutte le misure μ' invarianti rispetto alla trasformazioni S_t e tali che $\mu'(A) = 1$, la μ è quella che *uguaglia* il massimo della espressione:

$$\max_{\mu'} \{ \frac{1}{t} s(\mu') - \frac{1}{t} \int_A \mu'(dx) \log(\det J_t(x)) \} \tag{5}$$

ove $s(\mu')$ è l' entropia del sistema dinamico (A, S_t, μ'), (v. Entropia e informazione) e $J_t(x)$ è la matrice jacobiana della trasformazione S_t pensata come trasformazione di A in se (Bowen–Ruelle, 1975). Quest' ultima nozione ha evidentemente senso solo se A è una superficie regolare. Poichè gli attrattori spesso non sono superfici regolari è necessario estendere la nozione di piano tangente ad A in modo che $J_t(x)$ abbia senso. [Si deve, allo scopo, supporre che A sia pensabile come unione di superfici, regolari nelle vicinanze di ogni loro punto x e di uguale dimensionalità al variare di x, sulle quali la trasformazione S agisce in modo "espansivo", ossia asintoticamente dilata gli elementi di linea tangenti alla superficie con esponenti di Lyapunov > 0, v.; e se V_x^i è un elemento della superficie regolare tangente ad A che passa per x dovrà inoltre essere vero che $V_x^i \subset A$, e che $S_t V_x^i \cap V_{S_t x}^i$ è un elemento della superficie regolare in $S_t x$, di uguale dimensionalità. In tale situazione si può evidentemente definire $J_t(x)$; la superficie V_x^i si dice *(parte locale della) varietà instabile* per x, o anche *superficie dell' attrattore A*. Si pensa che gli attrattori strani abbiano le properità geometriche necessarie a che questa definizione abbia senso: tali proprietà vanno sotto il, nome di *proprietà di iperbolicità*. Si veda il capitolo Ergodicità e Irreversibilità]. Poichè si può mostrare che $s(\mu')$ è proporzionale a t e che l' altro termine in (5) è indipendente da t, il valore di t non ha importanza in (5), purchè $t > 0$.

L' analogia con la MSC, osservata da Ruelle e espressa dalla (5), in cui $\log J_t(x)$ gioca il ruolo di densità di energia potenziale, è talmente elegante che si tende a richiedere, nel caso dei sistemi dissipativi, che la (5) venga aggiunta alle proprietà fondamentali (cfr. (4) e righe seguenti) definenti gli attrattori strani. Qui non seguiremo questa strada, tuttavia un attrattore strano A nel senso qui adottato che, inoltre, è tale che la sua statistica μ rende massima l' espressione (5) (e quindi A è tale che esistano le varietà instabili V_x^i, che sono necessarie per dare senso alla (5) stessa) sarà chiamato un attrattore strano con la proprietà SRB, e la misura μ si dirà una una

misura SRB, da J. Sinai, D. Ruelle, R. Bowen.

§7 Gli scenarii per lo sviluppo del caos.

L‘ idea della descrivibilità dei moti caotici dissipativi mediante attrattori strani, e dei moti caotici conservativi mediante insiemi invarianti con proprietà analoghe nel senso discusso, ha dato luogo ad un fecondo periodo di ricerca sulla teoria dei sistemi dinamici.

Di particolare interesse sono i meccanismi per cui, al variare di un parametro R una data equazione che descrive il sistema può passare da una situazione in cui i moti asintotici sono descritti da attrattori (caso dissipativo) o insiemi invarianti (caso conservativo) banali, (quali stati di equilibrio o moti quasi periodici a due o più frequenze), a una situazione in cui i moti sono descritti da attrattori strani o insiemi invarianti con moto caotico.

Considerando solo casi dissipativi il caos si sviluppa, al variare di un parametro R, diremo al crescere di R per fisssare le idee, secondo vari meccanismi, detti *scenarii* per l‘ inizio del moto caotico. Nei sistemi dissipativi i più comuni, che discutiamo a titolo illustrativo del concetto, sono:

1) scenario di Feigenbaum
2) scenario di Ruelle Takens
3) scenario dell‘ intermittenza

Il primo scenario ha inizio con un regime asintotico descritto da un valore di R in corrispondenza del quale il moto è asintoticamente periodico, cioè è regolato da un‘ orbita periodica attrattiva (quindi stabile). Al crescere di R, poi, l‘ orbita perde stabilità, per $R = R_1$, ed è sostituita, nel ruolo di attrattore, da un‘ altra orbita pure periodica ma di periodo (per $R = R_1$) doppio. Successivamente anche la nuova orbita perde stabilità, per $R = R_2$, ed è sostituita da una nuova orbita di periodo (per $R = R_2$) doppio, *etc.* Si ottiene così una successione di valori critici $R_1 < R_2 < R_3 \ldots$ del parametro di controllo in corrispondenza dei quali l‘ attrattore che descrive il moto, pur restando sempre un‘ orbita periodica e quindi "semplice", si va complicando sempre più, perchè l‘ orbita acquista periodo sempre più lungo, (variando poco fra un valore critico ed il successivo ma di un fattore 2 ad ogni valore critico). I valori R_n si accumulano, per $R \to \infty$, su un valore R_∞ e per $R \gtrsim R_\infty$ il comportamento asintotico può essere descritto da un attrattore strano (il che *non* significa che questo debba avvenire per tutti i valori $R > R_\infty$: di solito esisteranno, e saranno osservabili in esperimenti accurati, intervalli aperti nella regione $R > R_\infty$ nei quali il moto asintotico è descritto ancora da attrattori non strani).

Dal punto di vista della teoria della stabilità lo scenario in questione si può presentare quando le perdite di stabilità delle orbite periodiche avvengono perchè uno degli autovalori della matrice di stabilità passa per -1: se

questo accade un certo numero di volte è sempre più probabile che avvenga ancora un numero infinito di volte dando così luogo alla descritta *cascata* di Feigenbaum completa. Lo scenario è stato investigato in particolare da M. Feigenbaum (1976), che scoprì che una transizione al caos di questo tipo aveva importanti proprietà di universalità. Indipendentemente dal modello considerato, che sia un modello di evoluzione di popolazioni, o di reazioni chimiche o di moto di un fluido viscoso, Feigenbaum mostrò ad esempio che gli intervalli fra due valori critici successivi tendono a zero esponenzialmente con una ragione δ:

$$\frac{R_{n+1} - R_n}{R_n - R_{n-1}} \xrightarrow[n\to\infty]{} \delta = \delta_F = 4.66920.. \tag{6}$$

ove δ *non* dipende dal modello. Questa universalità va intesa nello stesso senso in cui si parla di universalità degli esponenti critici (v.) nelle transizioni di fase: pur essendo possibile trovare esempi matematici in cui δ esiste ma è diverso da δ_F, tuttavia $\delta = \delta_F$ è il valore "normale" che ci si aspetta di trovare in modelli che non godano di proprietà particolari. Ad esempio nei sistemi conservativi, che godono della particolare proprietà di essere descritti da equazioni hamiltoniane, c' è un analogo scenario e ivi il valore di δ è diverso da δ_F, pur essendo, nello stesso senso, universale nella classe dei sistemi conservativi.

Lo scenario di Ruelle Takens, invece, riguarda una situazione iniziale in cui il moto asintotico è descritto da un attrattore su cui ha luogo un moto quasi periodico a due frequenze (che quindi può essere pensato come svolgentesi su un prodotto di due cerchi, ossia su un toro bidimensionale). Al crescere del parametro di controllo R, per $R = R_c$, questo moto quasi periodico perde stabilità e genera, non già un attrattore quasi periodico a tre frequenze (come vorrebbe uno scenario basato sui moti quasi periodici) ma direttamente un attrattore strano.

Dal punto di vista della teoria della stabilità questo corrisponde al passaggio di una coppia di autovalori della matrice di stabilità, del toro 2-dimensionale, attraverso il cerchio unitario in punti complessi coniugati (la cui parte immaginaria potrebbe essere ingenuamente ma erroneamente interpretata come un terza frequenza che entra nel problema e, quindi, come il segno della nascita di un moto quasi periodico con tre frequenze). Le proprietà di universalità di una transizione al caos di questo tipo sono meno note che nel caso dello scenario di Feigenbaum e sembrano dipendere esplicitamente dalle proprietà diofantine del rapporto fra i due periodi del moto quasi periodico, (v. Teoria delle Pertubazioni).

Lo scenario dell' intermittenza invece prevede che l' attrattore sia, inizialmente, un' orbita periodica che, al crescere del parametro di controllo R, si avvicina (come curva nello spazio delle fasi) ad una simile orbita periodica

instabile, fino a confondersi con essa, per $R = R_c$, (si dice che ha luogo una *collisione* fra orbite periodiche di opposta stabilità) per poi, per $R \gtrsim R_c$, sparire (insieme con quella instabile con la quale si *annichila*).

Se O_c è la curva descritta dall' orbita periodica stabile, al valore $R = R_c$ corrispondente alla sua collisione con l' analoga orbita instabile, resta per $R \gtrsim R_c$ una traccia dell' orbita periodica O_c che esisteva per $R = R_c$: nel senso che un dato che si viene a trovare nelle vicinanze di O_c (che per $R > R_c$ non è più una traiettoria del sistema) tende a rimanervi a lungo seguendo da vicino la curva O_c (è questa una banale conseguenza dei teoremi di regolarità per le equazioni differenziali), tanto più a lungo quanto più R è vicino a R_c. Poi si allontana spendendo altrove un intervallo di tempo più o meno lungo per poi ritornare nelle vicinanze di O_c e ripetere una evoluzione simile. Si osserva dunque un moto che ha l' apparenza "regolare" (periodica) per un tempo abbastanza lungo (*fase laminare*), per poi cambiare carattere qualitativo per un intervallo di tempo di lunghezza casuale (attorno ad un certo valor medio) e poi tornare in una fase laminare, e così via indefinitamente. Si ha dunque un regime intermittente in cui si alternano fasi laminari e fasi non laminari: la durata media della fase laminare è, universalmente, dell' ordine $(R - R_c)^{-1/2}$, per $R \to R_c$.

Dal punto di vista della teoria della biforcazione questa collisione fra orbite (e conseguente annichilazione) corrisponde ad una perdita di stabilità dovuta al raggiungimento del valore $+1$ da parte di uno degli autovalori della matrice di stabilità (simultaneo ad un raggiungimento di $+1$ da parte dell' autovalore più grande di 1 della matrice di stabilità dell' orbita instabile con la quale avviene la collisione).

Questi tre scenarii sono tra i più comuni e meglio studiati. Possono apparire simultaneamente in uno stesso modello in diverse regioni di variabilità del parametro di controllo, ovvero possono descrivere l' evoluzione di attrattori che coesistono (ovviamente in diverse regioni dello spazio delle fasi) in dati intervalli di R. Inoltre non esauriscono tutte le possibilità: ad esempio la transizione al caos prima descritta nel modello di Lorenz non rientra in uno di essi e fornisce un quarto scenario: si ha inizialmente un punto di equilibrio, (uno qualsiasi dei due della forma $x = y = \pm(b(r-1))^{1/2}$, $z = r - 1$), che per $R = R_c$ perde stabilità perchè due autovalori della matrice di stabilità passano attraverso l' asse immaginario con parte immaginaria non nulla; ma per $R > R_c$ non appare un attrattore periodico con periodo uguale a $2\pi/\omega$, se $\pm\omega$ è la parte immaginaria dei due autovalori (che potrebbe a sua volta essere il seme da cui si origina una successiva transizione caotica secondo uno dei tre scenarii discussi). Bensì appare direttamente un moto caotico: vari esempi di questo scenario sono noti.

Dal punto di vista della teoria della stabilità si può vedere che, nel caso del

modello di Lorenz, esiste, per $R \lesssim R_c$, una orbita periodica instabile che per $R \to R_c$ ha periodo che tende a $2\pi/\omega$ e si viene a confondere con il punto fisso che così perde stabilità pur rimanendo esistente per $R \gtrsim R_c$ senza più l' orbita periodica ad esso vicina.

Infine si possono immaginare altri scenarii per l' origine di moti caotici i quali pur avendo poca rilevanza in normali sistemi dissipativi, possono divenire importanti in sistemi che godono di particolari proprietà che rendono impossibili o improbabili gli scenarii finora discussi. E' il caso delle transizioni caotiche in sistemi conservativi, in cui la struttura canonica delle equazioni costringe a vedere in modo un po' diverso la teoria dei moti caotici. Ad esempio la matrice di stabilità di un' orbita periodica deve avere autovalori raggruppabili in coppie con prodotto 1, e il teorema di Liouville implica che questi sistemi non possono avere attrattori non banali; il concetto di attrattore non ha più interesse e la formulazione dei vari problemi va rivista come accennato sopra (§6), con conseguente rivoluzionamento degli scenarii.

§8 Caos in sistemi più generali. Qualche esempio.

La teoria dei moti caotici ha ricevuto grande impulso dai problemi della turbolenza nei fluidi, attraverso i lavori di Lorenz Ruelle Takens e dalla teoria dei moti di sistemi meccanici semplici, attraverso i lavori di Poincaré, Boltzmann, Birkhoff, Kolmogorov–Arnold–Moser, ma in nessun modo si confonde con essi.

In generale si considerano moti descritti da una trasformazione $x \to Sx$, ove S è generata da un' equazione differenziale (come trasformazione che da l' evoluzione, dopo un tempo prefissato, di un dato iniziale generico) o da qualche altro meccanismo, quale ad esempio un algoritmo costruito allo scopo di risolvere un' equazione algebrica $P(x) = 0$ per iterazioni successive, come il metodo di Newton:

$$Sx = \frac{xP'(x) - P(x)}{P'(x)} \tag{7}$$

ove x è un punto del piano complesso. Ovvero può accadere che sia interessante, per altri motivi, studiare le iterate di una data trasformazione: la $x \to Sx$ può essere un modello discreto di evoluzione di un sistema, quale una popolazione, come nel caso della trasformazione *logistica*, $x \to \alpha x(1-x)$, ($x \in [0,1]$ e α è un parametro). Le trasformazioni S che trasformano insiemi in insiemi con volume strettamente più piccolo si dicono dissipative; quelle che invece conservano il volume si dicono trasformazioni che *conservano l' area* e contengono come sottoclasse le trasformazioni *hamiltoniane* (per le quali S è una trasformazione canonica) ossia *conservative* in senso proprio.

Ci si pone il problema del comportamento asintotico delle iterate della trasformazione S e se, in certe circostanze, tale comportamento sia caotico. Nel senso che per dati iniziali scelti a caso, con una data distribuzione di probabilità, in un certo insieme U è possibile trovare un insieme A per cui vale la (4), e la proprietà successiva ad essa, con un μ opportuna (in generale, se il sistema non è dissipativo, non si richiederà che A sia un attrattore nè che U sia aperto, v. §6). Ci si domanda inoltre, quando la trasformazione S dipende da un parametro di controllo R, se la μ sia banale per certi valori di R e non banale per altri (ossia, ad esempio, se il sistema dinamico (A, S, μ) è isomorfo ad uno schema di Bernoulli o ha entropia positiva o altro).

Esempi interessanti di sistemi non dissipativi che presentano moti caotici sono molti sistemi descriventi interazioni fra corpi celesti. Un esempio classico è quello di un moto di un pianetino che si muove sul piano ortogonale a quello sul quale ruotano due stelle identiche su un' orbita ellittica di eccentricità $\varepsilon > 0$. Si immagina che il moto kepleriano delle due stelle non sia influenzato dal pianetino, e che il pianetino si muova sulla retta che passa per il baricentro (fisso) delle due stelle. L' equazione del moto per la coordinata z del pianetino è:

$$\frac{d^2}{dt^2} z = -\frac{2kz}{(z^2 + r_\varepsilon(t)^2)^{3/2}} \tag{8}$$

ove k è il prodotto fra la costante di gravitazione e la massa di una stella e $r_\varepsilon(t)$ è la distanza di una delle stelle dal baricentro, in funzione del tempo. Si può decidere di osservare gli intervalli di tempo $\dots, t_{-1}, t_0, t_1, \dots$ fra i successivi passaggi, con $\dot{z} > 0$ del pianetino per il piano delle stelle e considerare la successione $[t_i/T]$, delle parti intere delle misure di questi intervalli in unità del periodo T del moto delle stelle (che è la unità di tempo naturale). Si può dimostrare (K. Sitnikov, 1961) che, se $\varepsilon > 0$ è abbastanza piccolo, $\bar{s}$ abbastanza grande e se $\dots, s_{-1}, s_0, s_1$ è una prefissata, ma *arbitraria*, successione di interi $s_j \geq \bar{s}$, allora si possono trovare dati iniziali $z_0, \dot{z}_0$ tali che $[t_i/T] = s_i$ per tutti i valori di i. Dunque è chiaro che il sistema è dotato di moti coatici. Però l' insieme U dei dati iniziali cui segue un moto caotico potrebbe avere misura nulla nello spazio delle fasi, e anche supponendo (cosa non nota) che invece abbia misura positiva ci si attende che accada che la misura che determina la statistica di un dato iniziale scelto a caso in U (con la distribuzione di Liouville ristretta ad U) dipenda dal dato scelto (cosa che di solito non avviene nel caso di sistemi disipativi dotati di attrattori strani).

Un altro tipico esempio di moti caotici in sistemi non dissipativi è fornito dallo studio della *trasformazione standard*; è la trasformazione (canonica):

$$p' = p + \frac{\varepsilon}{2\pi} \sin \frac{q}{2\pi}, \qquad q' = q + p' \tag{9}$$

Si può fissare un numero irrazionale r e studiare un moto che si possa descrivere come:

$$S^n(p,q) \equiv (P_\varepsilon(nr), nr + Q_\varepsilon(nr)) \tag{10}$$

ove $P_\varepsilon, Q_\varepsilon$ sono opportune funzioni regolari periodiche con periodo 2π. Si può verificare, sperimentalmente (cioè a mezzo di esperimenti numerici) che se r è un irrazionale quadratico (ad esempio se $r = \sqrt{2}$, o $r = r_a = (\sqrt{5}-1)/2$) allora questo moto esiste se ε è abbastanza piccolo (ed è un moto quasi periodico a due frequenze il cui rapporto è r). Però esiste un valore ε_c in corrispondenza del quale P_ε e Q_ε cessano di esistere e per $\varepsilon \overset{>}{\sim} \varepsilon_c$ non esiste alcun dato iniziale dotato di un moto quasi periodico della forma (10), nell' intorno dell' ultimo insieme percorso dal moto in questione (cioè l' insieme delle coppie $(p,q) = (P_{\varepsilon_c}(x), x + Q_{\varepsilon_c}(x))$, $x \in [0, 2\pi]$).

§9 *Rappresentazioni grafiche del caos. I frattali. Universalità.*

Di grande interesse, talvolta anche dal punto di vista estetico, sono le immagini grafiche che si ottengono quando si cerca di rappresentare moti caotici.

Se, ad esempio, si cerca di rappresentare un attrattore strano di un sistema dissipativo, o la frontiera che separa i bacini di attrazione di attrattori diversi (che spesso è teatro di moti caotici, come nell' esempio ben noto della frontiera che separa i bacini di attrazione delle tre radici terze dell' unità nello studio delle iterate della trasformazione $Sx = x - (x^3-1)/3x^2$, generata dal metodo di Newton per la soluzione dell' equazione $x^3 - 1 = 0$) si trovano spesso insiemi *frattali*.

La caratteristica di un insieme frattale A è di essere dotato di proprietà di invarianza di scala. Empiricamente questa proprietà si manifesta al modo seguente: si fissa un rettangolo $\mathcal{R}_0$ di lati $l_1^0, \ldots, l_d^0$, se d è la dimensione dello spazio cartesiano in cui A è collocato, e si studia la intersezione $A_0 = A \cap \mathcal{R}_0$ di $\mathcal{R}_0$ con l' insieme. All' interno di $\mathcal{R}_0$ si identifica un altro rettangolo $\mathcal{R}_1$ di lati $l_1^1, \ldots, l_d^1$ più piccoli ma tale che all' interno di $\mathcal{R}_1$ l' insieme $A_1 = A \cap \mathcal{R}_1$ appaia avere la stessa forma di A_0 o assai vicina; e così si continua costruendo, se possibile, altri insiemi $\mathcal{R}_2, \mathcal{R}_3, \ldots$. Se per $n \to \infty$ le direzioni dei lati dei rettangoli tendono a valori limite e se i rapporti l_n^i / l_{n+1}^i, $i = 1, 2, \ldots, d$, tendono a valori limite $\lambda^i \geq 1$ e se gli insiemi $A_n = A \cap \mathcal{R}_n$, riscalati di un fattore $(\lambda^i)^n$ nella direzione dell' i-mo lato, tendono ad una forma limite A_∞, allora si dice che A è invariante di scala nell' intorno del punto $\cap_{n=0}^{\infty} \mathcal{R}_n$, con coefficienti di dilatazione $\lambda^1, \lambda^2, \ldots \lambda^d$ e forma limite A_∞. L' insieme A si dice *frattale* se per ogni punto x appartenente ad A si può eseguire una costruzione simile. Se A è un attrattore strano e se è un frattale nel senso ora descritto, si può dimostrare che i coefficienti di

dilatazione possono essere scelti in modo da non dipendere da $x \in A$, con probabilità 1 rispetto alla misura μ che descrive la statistica dei moti su A, cfr. (4).

Esempi semplici di insiemi frattali sono appunto la frontiera di separazione fra i bacini di attrazione delle tre radici dell' unità per la trasformazione del piano complesso generata dal metodo di Newton per l' equazione $x^3-1=0$; ovvero la frontiera che separa dagli altri punti del piano complesso l' insieme dei punti z che sotto iterazione della trasformazione $z \to z^2+\alpha$ non tendono all' ∞: per $\alpha \neq 0$ questo insieme è frattale non banale (si noti che insiemi A che sono superfici o linee regolari sono, ovviamente, frattali ma da questo punto di vista sono banali). Un ulteriore esempio di insieme frattale è l' attrattore che descrive il moto asintotico nel modello di Lorenz per valori $r \gtrsim r_c$.

Da quanto ora discusso emerge che si possono associare ai moti caotici grandezze che in qualche modo ne descrivono la caoticità (come ad esempio i coefficienti di dilatazione della eventuale struttura frattale di un attrattore strano). E' un fatto sperimentale che alcune proprietà quantitative dei moti caotici hanno "proprietà di universalità", ossia comuni a vaste classi di moti caotici, descritti da modelli anche apparentemente assai diversi. E' quindi naturale che nella ricerca si ponga un grande accento su queste proprietà di universalità, che consentono di mettere un po' d' ordine e di impostare una classificazione fra fenomeni che a prima vista appaiono assolutamente diversi.

Un esempio di proprietà universale è stato già illustrato nello scenario di Feigenbaum, in cui appare il numero universale $\delta_F = 4.66920..$, insieme ad altri numeri, pure universali che non abbiamo discusso e che descrivono altri aspetti della cascata di sdoppiamenti dei periodi dell' attrattore.

Un altro esempio di proprietà universale può essere trovato nella transizione caotica della trasformazione standard, prima introdotta (v. §8): le funzioni $P_\varepsilon, Q_\varepsilon$, che caratterizzano il moto in esame divengono singolari pper $\varepsilon \to \varepsilon_c$ e la natura delle singolarità dipende dal numero r, però se r viene mantenuto fisso e la trasformazione standard è sostituita da una trasformazione ad essa vicina la natura della singolarità non muta (così, per $\varepsilon = \varepsilon_c$, le funzioni $P_\varepsilon(x), Q_\varepsilon(x)$ pur restando continue non sono più differenziabili ma presentano, su un insieme denso di valori dell' argomento x una singolarità a potenza di ordine α, ossia se x_0 è uno dei punti dell' insieme delle singolarità la variazione delle funzioni $P_{\varepsilon_c}, Q_{\varepsilon_c}$ fra x_0 e un punto vicino x è dell' ordine di $|x-x_0|^\alpha$, con α che è *universale*: dipende

da r ma *non* dalla particolare trasformazione considerata).

§*10 Conclusioni e prospettive.*

In generale esiste una vasta fenomenologia sulla struttura dei moti caotici, e dal punto di vista teorico si usa discutere quali siano le possibilità che, *a priori*, ci si aspetta di incontrare: cioè attraverso quali scenarii è possibile o probabile che moti caotici si manifestino e quali possano essere le proprietà quantitative universali ad essi associate. L' analisi ha, sotto certi aspetti, un sorprendente grado di affidabilità al punto che è stato a volte detto, da autorevoli sperimentatori, che è inutile eseguire fino in fondo certi esperimenti tanto si può essere certi dei loro risultati. Ad esempio se nello studio dei moti di un sistema dissipativo appare che il sistema è descritto da un attrattore periodico che si sdoppia più volte, al variare di un parametro di controllo, è praticamente inutile investire tempo e strumenti per cercare di verificare più accuratamente se gli sdoppiamenti continuino in una cascata infinita e se i punti critici si accumulino con la progressione geometrica $\delta_F = 4.6692..$ prevista dallo scenario di Feigenbaum.

Però, al tempo stesso, la teoria è in uno stato molto insoddisfacente, in quanto non si conoscono metodi, altro che empirici e di dubbia generalità, per prevedere a priori se una certa evoluzione generi moti caotici e, se sì, con quale scenario ed in corrispondenza di quali valori dei parametri di controllo. Esiste ormai una grande messe di dati che si possono classificare sulla base di pochi scenarii, ma non si conosce una utile teoria unificatrice. I risultati sperimentali e numerici si riducono spesso ad aride tabulazioni dell' ordine in cui i vari scenarii si presentano al variare dei parametri di controllo, e la situazione ricorda quella dello studio delle linee spettrali atomiche prima della teoria di Bohr Sommerfeld e della equazione di Schrödinger: allora i lavori sperimentali producevano serie di righe spettrali, notavano regolarità, ma restava misterioso il perchè ci fossero e perchè si presentassero in un modo anzichè in un altro.

Come osservazione conclusiva è utile sottolineare che la teoria dei moti caotici ha anche contribuito a chiarire due punti che erano certamente dibattuti ancora intorno al 1970. Il primo è che non sono necessari infiniti (o anche solo molti) gradi di libertà per osservare fenomeni di moto disordinato, quali la turbolenza (*a fortiori*, dunque, anche fluidi bidimensionali, (che già hanno infiniti gradi di libertà) possono presentare fenomeni di turbolenza). Il secondo è che non è in alcun modo necessario un rumore di fondo, ossia una perturbazione stocastica agente sul sistema, perchè si possano osservare fenomeni di caos: i sistemi deterministici anche molto semplici, (come il sistema di Lorenz), possono presentare moti caotici e ricche fenomenologie di scenarii; al punto di poter essere usati come generatori

di numeri casuali (e proprio su questo tipo di sistemi sono basati i generatori di numeri casuali negli elaboratori elettronici). Ovviamente un rumore di fondo può rendere caotico un moto che non sarebbe altrimenti tale; ma in questo caso il disordine diminuirà al diminuire del rumore di fondo. Invece un rumore di fondo è senza effetto, almeno se piccolo, su un moto disordinato di un sistema deterministico: non appena il rumore diventa piccolo rispetto al caos intrinseco del sistema, la sua presenza diventa trascurabile, e questo può essere efficacemente sfruttato nell' interpretazione dei dati sperimentali e, per diminuirne i costi, nella progettazione di esperienze.

Bibliografia.

[1]J. Gleick: *Caos: la costruzione di una nuova scienza*, Rizzoli,1989.
[2]J. Eckmann: *Roads to turbolence in dissipative systems*, Review of Modern Physics, 53, 643-65via 4, 1981.
[3]J. Eckmann, D. Ruelle: *Ergodic theory of chaos and strange attractors*, Review of Modern Physics, 57, 617-656, 1985.
[4]P. Cvitanovic: *Universality in chaos*, Hilger, Bristol, 1984.
[5]D. Hofstadter, *Methamagical themes*, Scientific American, 245, 16-29, 1981; traduzione it. in Le Scienze, 1981.

Voci lessicali per la sezione di Meccanica Statistica Classica,

Giovanni Gallavotti,
voci per il Dizionario delle Scienze Fisiche
della Enciclopedia Italiana
Agosto 1989

Ammasso: (inglese: "cluster"): in teoria della percolazione; dato un reticolo e prefissato un insieme di coppie di punti del reticolo si dice che i punti delle coppie prefissate sono connessi o formano un legame: un ammasso è un insieme di punti del reticolo collegati da un cammino connesso di legami. In teoria dei gas rarefatti la stessa nozione appare nella classificazione dei diagrammi necessari alla costruzione dei coefficienti della serie del viriale per l' equazione di stato (cluster expansion).

Assiomi della geometria, (di Hilbert): sono gli assiomi sui quali si può fondare la geometria elementare (o "euclidea").
Dal punto di vista logico si dimostra che non contengono altre contraddizioni oltre quelle (eventualmente) presenti nell' aritmetica.
Gli assiomi stabiliscono le proprietà di relazioni fra enti astratti detti "elementi geometrici" (non definiti) che sono chiamati *punti, rette, piani*; le relazioni (non definite) di incidenza ("passa per"), ordine ("si trova tra"), congruenza ("è congruente a") hanno proprietà formulabili (ad esempio) come 20 assiomi divisi in cinque gruppi.
I) *assiomi di incidenza*: permettono di dar senso alla locuzione "passa per" (o "contiene" o sinonimi) e alle sue versioni passive ("appartiene a", o "giace su" o sinonimi)
1) esiste una retta che passa per due punti comunque dati.
2) per due punti distinti passa al più una retta.
3) ogni retta passa per almeno due punti ed esistono 3 punti che non sono su una stessa retta.
4) per tre punti per i quali non passa una retta ("non collineari") passa un piano ed ogni piano contiene almeno un punto.

5) per tre punti per i quali non passa una retta passa al più un solo piano.
6) se due punti sono su un piano allora la retta che li contiene giace sullo stesso piano.
7) se due piani hanno un punto in comune allora hanno in comune almeno una retta per quel punto.
8) esistono quattro punti che non giacciono su un piano.
Qui la locuzione "passa per" ha lo stesso significato di "contiene" e la locuzione "giace su" ha lo stesso significato di "è contenuto in". Se due rette hanno un punto in comune ("passano per lo stesso punto") si dice che si "intersecano" nel punto, ecc.
II) *Assiomi d' ordine*: permettono di stabilire le proprietà di una "relazione d' ordine" fra i punti di una retta, espressa dalla locuzione "si trova tra":
1) se un punto B si trova tra A e C, allora A, B, C sono due a due diversi e B si trova tra C ed A.
2) dati due punti A e C allora esiste, sulla retta che passa per A e C un punto B tale che C si trova tra A e B.
3) dati tre punti (due a due distinti) su una retta uno ed uno solo si trova tra gli altri due.
4) dati tre punti non collineari ("triangolo") ed una retta nel loro piano che taglia uno dei tre segmenti ("lati") delimitati dai tre punti allora essa taglia anche un altro dei tre segmenti.
Questi assiomi consentono di definire logicamente varie nozioni intuitive quali ad esempio il "segmento" (per la cui definizione l' ultimo assioma non è necessario), di punto "interno ad un segmento", di punto estremo di un segmento, di segmento interno ad un segmento, di "semiretta", di angolo fra due semirette uscenti da uno stesso punto, ecc.
III) *Assiomi di congruenza*: permettono di stabilire la nozione di "congruenza" (o uguaglianza o sovrapponibilità):
1) se A,B sono punti su una retta a e A' è su una retta a' allora su ogni semiretta di a' uscente da A' esiste un punto B' tale che il segmento AB è congruente a $A'B'$. Inoltre AB e BA sono congruenti.
2) se AB è congruente ad $A'B'$ e se $A'B'$ è congruente ad $A"B"$ allora AB è congruente a $A"B"$.
3) se AB e BC sono segmenti di una retta senza punti interni in comune e se anche $A'B'$ e $B'C'$ sono segmenti di una retta senza punti interni in comune, allora se AB è congruente a $A'B'$ e BC è congruente a $B'C'$ segue che AC è congruente a $A'C'$.
Gli assiomi precedenti consentono di definire la nozione di angolo e di poter quindi definire le proprietà della congruenza fra angoli, espresse dai seguenti due assiomi:
4) dato un angolo determinato da due semirette h, k uscenti da uno stesso punto e data una semiretta h' uscente da un punto A' e giacente su un

piano a', allora da ogni lato di a' esiste una semiretta k' tale che l' angolo hk è congruente all' angolo $h'k'$.
5) dati due triangoli ("terne di vertici non allineati") ABC e $A'B'C'$ tali che AB è congruente a $A'B'$, BC è congruente a $B'C'$ e l' angolo $A\hat{B}C$ è congruente a $A'\hat{B}'C'$, allora gli altri due angoli al vertice sono congruenti. Questi assiomi consentono, ad esempio, di impostare la teoria della ugua glianza dei triangoli.
IV) *Assiomi di continuità*: permettono, fra l' altro, di stabilire nozioni metriche quali la lunghezza di un segmento o l' ampiezza di un angolo.
1) Siano AB e CD due segmenti, allora sulla retta AB esistono n punti $A_0 \equiv A, A_1, \ldots, A_n$ tali che A_jA_{j+1} sono congruenti, per $j = 0, 1, \ldots, n-1$ a CD, A_j è tra A_{j-1} e A_{j+1} per $j = 1, \ldots, n-1$, e B è tra A e A_n.
2) dato un insieme di infiniti segmenti A_1B_1, A_2B_2, ... tali che A_iB_i contiene $A_{i+1}B_{i+1}$ per $i \geq 1$, e se per ogni fissato segmento CD si può trovare j tale che il segmento A_jB_j è più piccolo di CD, allora esiste un unico punto X comune a tutti i segmenti.
V) Infine, *dulcis in fundo*:
1) data una retta e un punto fuori di essa esiste al più una parallela alla retta data che passa per il il punto dato.
Il quarto assioma d' ordine è detto assioma di Pasch. Il primo assioma di continuità è detto assioma di Archimede. Il secondo assioma di continuità (detto assioma di Cantor) è sostituito, nella formulazione originale di Hilbert, dall' assiona di completezza:
"Gli elementi della geometria costituiscono un sistema che non può essere ampliato, mantenendo le relazioni stabilite dagli assiomi aggiungendovi nuovi elementi, cioè nuovi punti, rette o piani."
La geometria elementare nello spazio riposa dunque, nella formulazione ora descritta, su 20 assiomi; quella piana su 14 (poichè gli assiomi di incidenza $4 \div 8$ non si devono formulare). Ma evidentemente il numero degli assiomi dipende da come vengono formulati e di solito si pensa ad ogni gruppo come ad un singolo assioma e come costituenti "i cinque assiomi sui quali riposa la geometria euclidea": sono detti assiomi di Hilbert perchè da lui formulati.
La critica moderna tende a considerare come interpolazioni dei commentatori le strane definizioni di Euclide di punto retta e piano: è sostenibile che il punto di vista di Euclide fosse invece puramente assiomatico e essenzialmente coincidente con quello di Hilbert.
Bibliografia: N. Efimov: *Higher geometry*, MIR, Mosca, 1980; L. Russo: *Sulla non autenticità delle definizioni degli enti geometrici fondamentali contenute negli Elementi di Euclide*, Bollettino dei Classici, Accademia dei Lincei, 1993.

Assiomi dei numeri naturali, (di Peano): definiscono le proprietà dei numeri interi non negativi, o "numeri naturali" in termini di tre nozioni primitive (non definite): quella di numero naturale, quella di successore e quella di 0.
1) 0 è un numero naturale.
2) se x è un numero naturale, allora il successore di x, che si denota $x+1$, è un numero naturale. 3) non esiste un numero naturale del quale 0 sia il successore.
4) se x e y sono numeri naturali con successori uguali allora x e y sono uguali.
5) se per ogni numero naturale x è data una relazione logica $\varphi(x)$ e se $\varphi(0)$ è vera e la verità di $\varphi(x)$ implica quella di $\varphi(x+1)$ allora tutte le relazioni $\varphi(x)$ sono vere.
Bibliografia: P. Suppes, *Axiomatic set theory*, Dover, New York, 1972.

Attività: di una specie di molecole in un gas mono o multi molecolare è la funzione $z = e^{\beta\mu}(m\beta/2\pi)^{3/2}$ ove μ è il potenziale chimico, v..

Attrazione : dominio di: per un attrattore A delle soluzioni di una equazione differenziale o le iterazioni di una trasformazione. È l' insieme dei dati iniziali che evolvono in modo che la distanza da A tenda a 0.

Attrattore : per una equazione differenziale o per le iterazioni di una trasformazione. E' un insieme chiuso invariante A tale che dati iniziali abbastanza vicini ad A evolvono in modo che la loro distanza da A tenda a zero. Un attrattore puo' essere un punto fisso, un' orbita periodica o un oggetto più complesso. Se il moto sull' attrattore é caotico si dice che l' attrattore è strano, v. Caos.

Attrattore strano: v. Attrattore, Caos.

Autodiffusione, coefficiente di A.: in un gas o liquido in equilibrio termodinamico; è il valore limite D per $t \to \infty$ del rapporto fra il quadrato della distanza percorsa da una molecola o altra particella ed il tempo t in cui tale distanza è stata percorsa. Se il limite D non esiste o se vale 0 o ∞ si dice che il sistema presenta una diffusione anomala. In un gas rarefatto tridimensionale il moto delle particelle ha diffusione normale, mentre nel caso bidimensionale la diffusione è anomala (infatti si ritiene che il quadrato della distanza percorsa sia proporzionale a $t \log t$ invece che a t). Un esempio è il moto di atomi di idrogeno in un metallo che lo assorbe (ad esempio Pd), in cui a temperatura ambiente e pressione ambiente la costante di diffusione ha ordine di grandezza di qualche micron quadrato al secondo. Nel caso di particelle macroscopiche in sospensione colloidale il coefficiente di autodiffusione misura la rapidità del moto browniano ed è direttamente legato alla temperatura T e alle viscosità η e raggi a delle particelle in sospensione dalla relazione (di Einstein Smoluchovski): $D = 3k_BT/6\pi\eta a$ (v. Coefficienti di Trasporto). Questa relazione consente la misura di k_B

e quindi del numero di Avogadro a partire da misure macroscopiche sul moto browniano; è valida solo approssimativamente per il moto delle molecole perchè è basata sulla ipotesi che le particelle in sospensione siano talmente grandi da poter considerare il fluido un continuo.

Autoevitante: cammino aleatorio: è un cammino aleatorio svolgentesi su un reticolo (regolare o no) in modo da non intersecare se stesso; ossia ad ogni passo visita un punto diverso da tutti i precedenti.

Avogadro, numero di A.: Numero N_A di atomi contenuto in un volume $V_0 = 22.413\, litri$ di gas perfetto a temperatura $T_0 = 273.15^o K$ e a pressione $p_0 = 1.\, atm$. Il volume normale è per definizione quello occupato da $2\, g$ di idrogeno a temperatura T_0 e pressione p_0. Più in generale il numero di Avogadro è il numero di molecole contenute in una massa di composto omogeneo pari ad un numero di grammi uguale al peso molecolare. Il valore del numero di Avogadro è $N_A = 6.02\, 10^{23}$. La possibilità di definire il numero di Avogadro è basata sulla legge di Avogadro per i gas perfetti: volumi uguali di gas nelle medesime condizioni di temperatura e pressione contengono lo stesso numero di molecole. Nella legge dei gas perfetti la costante dei gas $R = 8.31 joule/^o K$ è proporzionale al numero di Avogadro, e il coefficiente di proporzionalità è la costante di Boltzmann: $R = k_B N_A$, v. MS. Il numero di Avogadro si può ricavare in vari modi da esperimenti macroscopici che consentono di calcolare la costante di Boltzmann. Classico è il metodo di Perrin. Particelle macroscopiche colloidali in una sospensione in equilibrio si distribuiscono con una densità $n(h)$ che dipende dall' altezza h nel contenitore; la densità varia, secondo la Meccanica Statistica, verificando la legge $n(h) = n(0) \exp(-mgh/k_B T)$ ove k_B è la costante di Boltzmann, g è l'accelerazione di gravità, T è la temperatura assoluta, m è la massa delle particelle in sospensione. Dalle misure si ricava k_B e quindi $N_A = R/k_B$. Il numero di Avogadro può essere ricavato con moltissimi altri metodi. Il metodo di Loschmidt lo deduce dalle relazioni $4\pi n \lambda a^2 \sqrt{2} = 1$ fra il cammino libero medio λ, la densità numerica n, il raggio a delle molecole (che consente di ottenere na^2 attraverso la misura di λ) e dalla relazione fra libero cammino medio, viscosità e temperatura, v. Cammino l Libero m Medio, v Viscosità, Autodiffusione, Moto Browniano) e dalla relazione $n\pi(2a)^3/6 = c(T)\rho_{gas}/\rho_{liquido}$, ove $\rho_{gas}, \rho_{liqido}$ denotano le densità del gas e del liquido alla temperatura T e $c(T)$ è una costante da determinare in funzione della temperatura T con argomenti teorici, ad esempio a partire dalla legge di van der Waals, v. serie del Viriale, MSC, (questa seconda relazione consente di misurare approssimativamente na^3). Da queste misure si deducono a e n e quindi $n_{moli} N_A = nV$. Einstein in una serie di lavori trovò ed esaminò una decina di metodi per il calcolo del numero di Avogadro. Celebri sono rimasti i metodi basati sul moto browniano (da cui deriva il metodo di Perrin) e quelli sulla opalescenza critica

o sul colore del

cielo. Una rassegna sul numero di Avogadro è in: Deslattes, Annual review of Physical Chemistry, 31, 435, 1980; si veda anche A. Pais, *Subtle is the Lord. The science and life of Albert Einstein*, Oxford University press, 1982, cap. II, §5.

Bernoulli, schema di: processo stocastico, v., i cui campioni sono successioni $\underline{\sigma} = (\ldots, \sigma_{-1}, \sigma_0, \sigma_1, \sigma_2, \ldots)$ di simboli $\sigma_i = 1, 2, \ldots, n$ che possono essere generate lanciando agli istanti $\ldots, -1, 0, 1, 2, \ldots$ un dado a n facce che hanno probabilità $p_1, p_2, \ldots, p_n$. Uno schema di Bernoulli a n simboli (o stati) è dunque determinato da n numeri $p_1, p_2, \ldots, p_n$ tali che $\sum_i p_i = 1$. Se due schemi di Bernoulli $(p_1, \ldots, p_n)$ e $(q_1, \ldots, q_n)$ hanno la stessa *entropia* S, ossia se $S \equiv -\sum_{i=1}^{n} p_i \log p_i = S' \equiv -\sum_{j=1}^{m} q_j \log q_j$, allora sono isomorfi. E' possibile, cioè, definire una funzione $\underline{\eta} = X(\underline{\sigma})$, detta *codice*, che pone in corrispondenza biunivoca i campioni $\underline{\sigma}$ prodotti con il primo schema di Bernoulli con quelli $\underline{\eta}$ prodotti dal secondo in modo che insiemi di campioni corrispondenti abbiano uguali probabilità di essere realizzati (teorema di D. Ornstein); la corrispondenza biunivoca ora menzionata viene stabilita però a patto di escludere dai due insiemi di campioni da porre in corrispondenza due insiemi a probabilità totale nulla (si dice che uno schema di Bernoulli è codificabile nell' altro "modulo 0").

Biliardo, di J. Sinai: v. Sinai.

Bogolubov, Born, Green, Kirkwood, Yvon, gerarchia di BBGKY: è una famiglia di equazioni per un sistema di N particelle classiche, in un volume V, equivalente alle equazioni di Newton. Nel caso di sistemi in equilibrio termodinamico diviene la famiglia di equazioni di equilibrio detta equazioni Kirkwood–Salzburg, (v.). E' utile perchè le equazioni legano fra loro funzioni di correlazione (v.) di vari ordini e hanno formalmente senso nel *limite termodinamico*, $N/V \to \rho$, $V \to \infty$. Le equazioni sono molto utili per tentare di estendere la teoria cinetica a densità più alte di quelle alle quali è applicabile l' equazione di Boltzmann. N. Bogolubov ha sviluppato un metodo per lo studio di queste equazioni assai analogo al metodo di Chapman Enskog per lo studio dell' equazione di Boltzmann. Come nel metodo di Chapman Enskog si distinguono tre stadi nell' evoluzione verso l' equilibrio e il metodo di Bogolubov consente di studiare la fase cinetica (che è fuori dall' analisi di Chapman Enskog) e la fase idrodinamica (v. metodo di Chapman Enskog) e mostra che nelle fasi in questione si può utilizzare, per lo studio della dinamica, un' equazione che estende l' equazione di Boltzmann. Formalmente il metodo di Bogolubov conduce ad un algoritmo che produce correzioni sistematiche all' equazione di Boltzmann interpretabili come uno sviluppo in serie del parametro na^3 ove n è la densità numerica e a il raggio della sezione d' urto. Tuttavia il metodo soffre di una grave difficoltà, la cui soluzione non è ancora veramente nota: il modo più semplice di illustrare la difficoltà è di analizzare le previsioni

della teo-

ria per i valori dei coefficienti di trasporto, quali il coefficiente di diffusione o la viscosità. Scrivendo tali quantità come il valore che avrebbero se la teoria dei gas rarefatti fosse applicabile (cosiddetto *valore di Boltzmann*) moltiplicato per una serie in $\varepsilon = na^3$ (con primo termine 1), si trova, come scoperto da E. Cohen (1961), che i coefficienti delle serie risultano essere divergenti a partire dal secondo.
La generalizzazione delle equazioni BBGKY a sistemi quantici è possibile e viene di solito considerata tale la equazione KMS (v.).

Bohr–Sommerfeld, quantizzazione: è la regola di quantizzazione in uso prima dell' introduzione dell' equazione di Schrödinger. A differenza di quest' ultima è applicabile solo a sistemi meccanici classicamente integrabili (v. Teoria delle perturbazioni). Se $(\underline{A}, \underline{\varphi})$ sono le variabili di azione angolo per un sistema integrabile a l gradi di libertà e se $E(A_1, \ldots, A_l)$ è la energia espressa in queste variabili, la regola di B.S. vuole che i livelli energetici possibili siano solo quelli per cui $E = E_{n_1,\ldots,n_l} = E(n_1\hbar, \ldots, n_l\hbar$ ove $\hbar$ è la costante di Planck divisa per 2π ($\hbar = 1.054\,10^{-27} erg\,sec$) e n_i sono numeri interi. La regola impone, dunque, che i soli livelli possibili siano quelli per cui le variabili di azione sono multipli interi di $\hbar$. Si dimostra che tale regola fornisce una prima approssimazione ai valori dei livelli calcolati via l' equazione di Schrödinger; le correzioni sono proporzionali a $\hbar$.

Boltzmann , costante di B.: rapporto k_B fra variazione di entropia e variazione del logaritmo del numero di stati microscopici che corrispondono agli stessi stati macroscopici di equilibrio termodinamico. E' anche il rapporto fra la costante dei gas R ed il numero di Avogadro (v. Avogadro) N: $k_B = R/N = 1.3805\,10^{-16} erg\,{}^oK^{-1}$. La energia cinetica media di traslazione di una molecola di un gas perfetto a temperatura assoluta T è $3k_BT/2$, (v. MSC, Equipartizione).

Boltzmann, ipotesi ergodica di B.: si immagina di dividere lo spazio delle fasi, di un sistema meccanico con N particelle, in cellette microscopiche di volume h^{3N}: in tal modo ogni celletta rappresenta una configurazione microscopica possibile delle N particelle. Si osserva l' evoluzione nel tempo di ciascuna configurazione osservandola ad intervalli di tempo multipli di un prefissato intervallo di tempo microscopico (tipicamente dell' ordine di $10^{-12}\,sec$). L' i.e. suppone che tale evoluzione consista in una permutazione delle cellette e che tutte le cellette di uguale energia facciano parte dello stesso ciclo di permutazione. In altre parole si suppone che ogni celletta di data energia visiti successivamente tutte le altre di uguale energia. Il tempo necessario perchè una configurazione iniziale assuma successivamente tutte le altre forme di uguale energia è il tempo di *ricorrenza* microscopico e fu stimato da Boltzmann stesso come proporzionale alla unità di tempo microscopico tramite un fattore di ordine $N^{N/2}$, v. MSC, che rende il tempo di ricorrenza di gran lunga superiore all' età dell' universo non appena

N supera qualche decina. Per l‘ etimologia di ergodico si veda v. Monodo.

Boltzmann, statistica di B.: in Meccanica Statistica quantica di N particelle identiche è la statistica in cui non si suppone che la funzione d‘ onda sia simmetrica o antisimmetrica per permutazioni delle particelle. Questa statistica ha solo un interesse teorico in quanto le particelle reali verificano le statistiche di Fermi-Dirac o di Bose-Einstein, v..

Boltzmann-Grad, limite di B.G.: situazione in cui si viene a trovare un gas rarefatto di molecole con sezione d‘ urto a^2 (ossia di raggio circa a) e densità n quando $a \to 0$, $n \to \infty$ in modo che $\lambda = (na^2)^{-1}$, e cioè il *libero cammino medio* λ, resti fisso. In questa situazione limite si ritiene che la equazione di Boltzmann descriva esattamente l‘ evoluzione di uno stato iniziale caotico (verificante cioè l‘ ipotesi che la funzione di distribuzione a n particelle è prodotto di n funzioni di distribuzione a 1 particella (caos molecolare)): lo stato evolverebbe restando caotico nello stesso senso e inoltre la funzione di distribuzione a 1 particella verificherebbe l‘ equazione di B.. Questa proprietà è stata dimostrata rigorosamente solo nel caso particolare di un sistema di sfere rigide e per tempi brevi rispetto al *tempo di volo*,ossia al tempo medio di percorrenza di una distanza pari al cammino libero medio λ. (v. MSC, Teorie Cinetiche).

Bose, condensazione di B.: transizione di fase in un gas perfetto verificante la statistica di Bose Einstein, ossia tale che le funzioni d‘ onda che descrivono gli stati delle N particelle sono simmetriche. Questo sistema, a temperatura abbastanza bassa e densità ρ maggiore di un certo valore critico $\rho_c(T)$, si presenta in uno stato di equilibrio in cui una frazione macroscopica del numero totale di particelle si trova nello stato di energia minima. Poichè lo stato di energia minima è descritto da una funzione d‘ onda costante su tutto il volume occupato dal sistema, si trovano $N_c = N(\rho - \rho_c)/\rho$ particelle (con $0 < \rho_c < \rho$) con funzione d‘ onda uguale e uniforme su tutto il volume (cioè con impulso nullo). Il fenomeno è particolarmente interessante perchè la separazione delle fasi avviene nello spazio degli impulsi e non, come nelle transizioni di fase classiche, nello spazio delle coordinate (con conseguente separazione fisica delle due fasi): il sistema resta perfettamente omogeneo nello spazio ambiente. Non è mai stato dimostrato che sistemi di particelle interagenti possano presentare una transizione di fase di questo tipo o se il fenomeno venga necessariamente smussato in presenza di interazione: si pensa che il fenomeno permanga in presenza di interazione e varie teorie lo considerano. (v. Statistiche quantiche, MSC).

Bose–Einstein, statistica di B.E.: è verificata da un sistema di N particelle identiche se gli stati quantici in cui tali particelle possono trovarsi sono descritti da funzioni d‘ onda simmetriche nelle coordinate delle particelle.

Bosone: nome che designa una particella che insieme ad altre particelle

ad essa identiche verifica la statistica di Bose–Einstein (v.).

Burnett, equazioni di B.: equazioni per il moto di un fluido classico ottenute ricercando le correzioni alle equazioni di Navier–Stokes (v.). Si deducono a partire dall' equazione di Boltzmann spingendo al secondo ordine le soluzioni formali date dallo sviluppo di Chapman–Enskog (v. metodo di Chapman Enskog, Teorie cinetiche, Coefficienti di trasporto). Trovano applicazione, nella forma linearizzata, nella teoria della dipersione ed assorbimento del suono in un gas rarefatto a lunghezze d' onda dell' ordine del libero cammino medio, ma la loro correttezza è in realtà una questione che viene spesso sollevata.

Cammino libero medio, distanza media percorsa da una molecola di gas o liquido fra due successive collisioni. L' espressione del libero cammino medio, nei gas rarefatti, monomolecolari, interagenti con forze a corta portata) è (Maxwell): $\lambda = 1/4n\pi a^2\sqrt{2}$, ove $4\pi a^2$ è la sezione d' urto totale di collisione (cioè a è una misura del raggio delle molecole) e n è la densità numerica. Una semplice relazione (Clausius–Maxwell) lega cammino libero medio λ, densità $\rho = nm$ (ove m è la massa delle molecole), velocità media v e viscosità η in un gas rarefatto: $\eta = \rho\lambda v/3$.

Campo medio, teoria del C.M. nelle transizioni di fase: approssimazione usata in MS per ridurre un problema a molti corpi ad un problema ad un solo corpo. Si suppone che l' interazione di un corpo con gli altri sia uguale al suo valore medio, che viene considerato come parametro. In tal modo l' interazione di ogni corpo con gli altri viene sostituita da una interazione con un campo esterno, che viene lasciato come parametro incognito. E' allora di solito (anche se non sempre) facile calcolare le proprietà termodinamiche del sistema che ora appare come un sistema di corpi indipendenti posti in un campo esterno. Le proprietà statistiche sono funzioni del valore assunto dal c.m. (parametro incognito) e, in particolare, il valore stesso del c.m.: questo dà, quindi, luogo ad una equazione di consistenza che determina il valore del c.m.. Se questa equazione ha più di una soluzione si interpretano le varie soluzioni come descriventi diversi stati di equilibrio possibili e quindi come manifestazione dell' esistenza di una transizione di fase. Le più semplici teorie di c.m. conducono alla teoria di van der Vaals per le transizioni liquido–gas, alla teoria di Curie–Weiss per il ferromagnetismo, al modello BCS per la superconduttività, che sono fra gli esempi più noti di teorie di c.m.. Le teorie di c.m. si rivelano di solito inadeguate nelle vicinanze del punto critico perchè prevedono singolarità diverse da quelle osservate (v. Esponenti Critici). A volte prevedono transizioni di fase in sistemi che si può mostrare *a priori* esserne, invece, privi (ad esempio in sistemi unidimensionali con forze a corta portata). (v. Transizioni di Fase; Serie e Teorema del Viriale). Esiste, tuttavia una situazione limite in cui la teoria di campo medio è esatta (v. Potenziale di Kac).

Canonico, insieme: (v. Distribuzione Canonica, Insiemi Statistici, MSC).

Celletta, dello spazio delle fasi: è in MSC una regione dello spazio delle fasi (per N particelle, e quindi $6N$ dimensionale) di larghezza Δp in ogni coordinata di impulso e Δq in ogni coordinata di posizione. I valori Δp e Δq sono i valori della precisione massima con cui si suppone di poter misurare le coordinate in questione. Nella MSC si scelgono Δp e Δq come parametri arbitrari che prima o poi vengono fatti tendere a zero. A volte però si scelgono Δp e Δq in modo che $\Delta p \Delta q = h$ ove h è la costante di Planck: questo per tenere in qualche modo conto del principio di indeterminazione. Se in corrispondenza di certi valori dei parametri termodinamici i risultati dei calcoli teorici dipendono dal valore di h in modo sensibile, al variare di h nell' intorno del valore della costante di Planck, se ne deduce che ci si trova in un regime in cui non si può più usare la Meccanica Statistica Classica e il problema va trattato usando la Meccanica Statistica Quantica. (v. MSC, IS, Statistiche quantiche).

Chapman Enskog: metodo di Chapman, I., Enskog, D.,metodo per la soluzione dell' equazione di Boltzmann e la sua connessione con la gasdinamica macroscopica, cioè con le equazioni di Eulero o Navier Stokes comprimibili. L' evoluzione verso l' equilibrio di un gas rarefatto, monomolecolare e con interazione a corta portata, inizialmente in uno stato di non equilibrio, attraversa tre stadii la cui durata dipende da tre scale di tempo associate a tre lunghezze caratteristiche. La portata del potenziale di interazione a, il libero cammino medio λ e una lunghezza L macroscopica, ad esempio una dimensione del contenitore. Le tre scale di tempo t_0, t_{lcm}, t_{macr} si ottengono dividendo le tre lunghezze per una velocità molecolare tipica (le prime due), come la velocità media, o per una velocità macroscopica tipica (la terza), come quella del suono. Nel caso di $1cm^3$ di idrogeno a 0^oC e $1atm$le tre lunghezze sono rispettivamente dell' ordine di $10^{-8}cm$, $10^{-5}cm$, $1cm$ e i tempi corrispondenti dell' ordine di $10^{-12}s$, $10^{-9}s$, $10^{-4}s$. Dunque le tre scale di tempo sono ben diverse. Per tempi $t \leq t_0$ il gas evolve in modo complicato e, se lo stato iniziale è lontano da uno stato privo di correlazioni e localmente maxwelliano, è necessario descriverne l' evoluzione usando le equazioni di Newton (ad esempio nella forma detta BBGKY); segue per $t_0 \ll t$
$lllt_{lcm}$ una evoluzione che dovrebbe essere descritta dall' equazione di Boltzmann, in cui le funzioni di correlazione a più particelle sono fattorizzate, cosi' che la funzione di distribuzione $f(\underline{r}, \underline{v}, t)$ che appare nell' equazione di Boltzmann fornisce una descrizione completa del gas; segue lo stadio *idrodinamico*, $t \gg t_{macr}$, in cui la funzione di distribuzione a un punto viene a dipendere dal tempo solo attraverso le tre funzioni $n(\underline{r}), T(\underline{r}), \underline{u}(\underline{r})$ che definiscono la densità media in $\underline{r}$, la temperatura media in $\underline{r}$ e la velocità media in $\underline{r}$. Dunque C.E. suppongono che durante lo stadio idrodinamico

f si possa scrivere come $f(\underline{r},\underline{v},t) = F(\underline{r},\underline{v};n(\underline{r}),T(\underline{r}),\underline{u}(\underline{r}))$ con la dipendenza dal tempo t che si manifesta solo perchè $n,T,\underline{u}$ dipendono da t. La discussione euristica di questa ipotesi suggerisce che essa possa essere valida solo se le grandezze $n,T,\underline{u}$ variano poco su una distanza dell' ordine del libero cammino medio, che equivale all' ipotesi che il parametro $\mu = t_{lcm}/t_{macr}$ sia molto minore di 1. Fatta questa ipotesi si immagina che F ammetta uno sviluppo della forma:

$$F(\underline{r},\underline{v};n,T,\underline{u}) = F_0(\underline{r},\underline{v};n,T,\underline{u}) + \mu F_1(\underline{r},\underline{v};n,T,\underline{u}) + \dots \tag{1}$$

e che le derivate spaziali dell F_j siano dell' ordine di grandezza di μ. Si scrive quindi l' equazione di Boltzmann come:

$$\frac{\partial F}{\partial t} = -\underline{v}\cdot\frac{\partial F}{\partial \underline{r}} + Q(F,F) \tag{2}$$

ove Q è il termine di collisione (v. MSC, Teorie cinetiche) e si sviluppa il membro di destra in una serie in cui i termini dello stesso ordine di grandezza in μ vengono raccolti insieme. Lo scopo è quello di fare lo stesso per il termine di sinistra e di ottenere cosi' una famiglia di equazioni che impongono la validità dell' E. di B. ad ogni ordine in μ. Poichè l' ipotesi fondamentale del metodo richiede che tutta la dipendenza dal tempo si manifesti attraverso le funzioni $n,T,\underline{u}$ occorre trovare uno sviluppo per le derivate temporali di queste grandezze. Questo viene ottenuto ricordando che, per definizione:

$$\begin{aligned} n(\underline{r}) &= \int d\,\underline{v} f(\underline{r},\,\underline{v}) \\ n(\underline{r})\,\underline{u}(\underline{r}) &= \int d\,\underline{v}\,\underline{v} f(\underline{r},\,\underline{v}) \\ n(\underline{r})\frac{3}{2}k_B T(\underline{r}) &= \int d\,\underline{v}\frac{m(\underline{v}-\underline{u}(\underline{r}))^2}{2} f(\underline{r},\,\underline{v}) \end{aligned} \tag{3}$$

e scrivendo le relazioni che si ottengono dall' E. di B. moltiplicandola membro a membro per $1, m\,\underline{v}, m(\underline{v}-\underline{u}(\underline{r}))^2/2$ e integrando rispetto a $\underline{v}$. Il punto chiave è che il termine di collisione a secondo membro non contribuisce (v. MSC, Teorie cinetiche) e allora si ottengono, rispettivamente:

$$\begin{aligned} \frac{\partial n}{\partial t} &= -\sum_{i=1}^{3}\frac{\partial n\,\underline{u}_i}{\partial\,\underline{r}_i} \\ \frac{\partial nm\,\underline{u}_j}{\partial t} &= -\sum_{i=1}^{3}[\frac{\partial nm\,\underline{u}_i\,\underline{u}_j}{\partial\,\underline{r}_i} - \frac{\partial \mathcal{T}_{ij}}{\partial\,\underline{r}_j})] \\ \frac{\frac{3}{2}\partial nk_B T}{\partial t} &= -\sum_{i=1}^{3}[\frac{\partial nk_B T(\underline{r})}{\partial\,\underline{r}_i} - \frac{\partial \mathcal{E}_i}{\partial\,\underline{r}_i} \end{aligned} \tag{4}$$

ove:

$$\begin{aligned}\mathcal{T}_{ij} &= \int m(\underline{v}_i - \underline{u}_i)(\underline{v}_j - \underline{u}_j) f(\underline{r}, \underline{v}) d\underline{v} \\ \mathcal{E}_i &= \int \frac{m(\underline{v} - \underline{u})^2}{2} (\underline{v}_i - \underline{u}_i) f(\underline{r}, \underline{v}) d\underline{v}\end{aligned} \tag{5}$$

Si riconoscono nelle equazioni (4,5) l' equazione di continuità della massa (v.), la prima equazione cardinale e la conservazione dell' energia, se $\mathcal{T}$ è interpretato come il tensore degli sforzi (v.) e $\vec{\mathcal{E}}$ è interpretato come flusso di calore (v.). Si sostituisce ora in (5) lo sviluppo (1) e si ottiene cosi' lo sviluppo in serie di μ del primo membro dell E. di B.: si impone quindi la validità dell' equazione di Boltzmann a *tutti gli ordini in* μ, tenendo conto nel conteggio dell' ordine in μ, nei due membri, che le derivate spaziali delle F_j vanno considerate di ordine μ. Si ottiene quindi una famiglia di equazioni per le F. Quella di ordine 0 impone che F_0 sia una maxwelliana (*equilibrio locale*):

$$F_0(\underline{r}, \underline{v}; n, T, \underline{u}) = n(\underline{r}) \frac{e^{-(\underline{v} - \underline{u}(\underline{r}))^2 m k_B T(\underline{r})/2}}{(2\pi/m k_B T(\underline{r}))^{3/2}} \tag{6}$$

Ottenuta questa relazione fondamentale si può risolvere l' equazione lineare per F_1 e poi per F_2 etc.. Si trova che le equazioni che si ottengono sono tutte risolubili se si impone la condizione di consistenza

$$\int F_j d\underline{v} \equiv 0, \int \underline{v}_i F_j d\underline{v} \equiv 0, \int (\underline{v}_i - \underline{u}_i)^2 F_j d\underline{v} \equiv 0, \quad j \geq 1 \tag{7}$$

e se si impongono alle funzioni F_j di verificare opportune equazioni (ottenute appunto imponendo la validità dell E. di B. ad ogni ordine). Si può poi vedere quali siano le equazioni (4,5) quando si trascurano le F_j con $j \geq j_0$. Si trovano, ed è questo il risultato fondamentale della teoria di C.E., le equazioni di Eulero comprimibili (per $j_0 = 0$), le equazioni di Navier Stokes comprimibili (per $j_0 = 1$) e altre equazioni che dovrbbero fornire correzioni alle equazioni di Navier Stokes (dette, per $j_0 = 3$ equazioni di Burnett). Nel caso delle equazioni di Navier Stokes si trova che corrispondono ad un tensore degli sforzi dato da:

$$T_{ij} = -\eta(\partial_j u_i + \partial_i u_j)$$

e ad una legge di trasporto del calore di Fourier:

$$\mathcal{E}_i = -\chi \partial_i T$$

con espressioni esplicite per η, χ, in termini del potenziale di interazione. Il metodo di C.E. è criticabile perchè il parametro μ in termini del quale lo

sviluppo è organizzato in realtà non appare nella equazione di B. e quindi la sua soluzione *ordine per ordine* non è chiarissima da un punto di vista matematico. Altri metodi per la teoria dell' equazione di B. cercano di evitare il problema. Fra questi il metodo di Hilbert che è in apparenza più rigoroso ma che soffre di problemi interpretativi; questo metodo conduce anche alle equazioni di Eulero e Navier Stokes però linearizzate, il che non è di per se un difetto in quanto la non linearità delle equazioni dedotte con il procedimentodi C.E. è illusoria perchè nelle ipotesi in cui ci si pone per la deduzione ci si attende che le equazioni linearizzate siano una buona approssimazione. L' ipotesi di bassa densità del gas è molto restrittiva: d' altra parte la equazione stessa di B. non può essere valida ad alta densità, per le ipotesi nelle quali viene dedotta, che implicano la trasurabilità delle collisioni multiple. Il più notevole fra i tentativi di estensione dell' equazione a densità più alte è di N. Bogoliubov: che però incontra gravi difficoltà di principio (v. gerarchia di BBGKY). Bibliografia: E.G.D. Cohen, in *Transport phenomena in fluids*, raccolto da H. Hanley, 1969, Dekker, New York.

Clausius Maxwell: Clausius Maxwell relazioni di C.M.: v. Cammino libero medio, Viscosità.

Conducibilità elettrica , (nei metalli): le teorie fenomenologiche più semplici sono le teorie di Drude, di Lorentz e di Sommerfeld. Le tre teorie suppongono che il metallo sia pensabile come contenente un gas di elettroni liberi e calcolano una relazione fra la conducibiltà elettrica σ_{el} in assenza di gradiente termico e la conducibilità termica σ_{term} in assenza di corrente elettrica. Questa relazione risulta in accordo con la legge di Wiedemann e Franz per cui, se k_B è la costante di Boltzmann, T la temperatura assoluta e e è la carica dell' elettrone si ha:

$$\sigma_{term}/\sigma_{el} = cost\,(k_B/e)^2\,T$$

e il valore della costante è 3 nel caso della teoria di Drude, 2 nel caso della teoria di Lorentz e $\pi^2/3$ nel caso della teoria di Sommerfeld.

Si intende qui per conducibilità termica quella dovuta ai soli elettroni di conduzione: ovviamente la conducibilià termica riceve anche contributi da parte del cristallo stesso, v. Teoria di Debye, distinguibili da quelli elettronici perchè si comportano diversamente per $T \to 0$ (proporzionali a T gli uni e indipendenti da T gli altri).

La caratteristica fondamentale delle tre teorie è l' ipotesi che il moto degli elettroni sia essenzialmente una successione di moti liberi su tratti di una lunghezza caratteristica l, interpretabile come libero cammino medio fra le collisioni degli eletttroni portatori di elettricità con gli atomi del cristallo che li contiene. Questi ultimi, a causa della agitazione termica, sono fuori dalle rispettive posizioni di equlibrio (le collisioni con gli atomi non sarebbero da considerarsi tali se essi fossero esattamente in equilibrio su un

cristallo regolare, perchè in tal caso non darebbero luogo a fenomeni dissipativi). Al termine di un cammino libero la velocità dell' elettrone cambia direzione con distribuzione uniforme (ed è qui che è importante che gli atomi del reticolo siano disordinati dall' agitazione termica: nel caso che fossero invece ordinati su un reticolo perfetto non si potrebbe supporre che le collisioni producano deviazioni casuali, chè invece sarebbero coerentissime).

La teoria di Drude suppone che il moto degli elettroni sia esattamente una spezzata di segmenti di lunghezza l; quella di Lorentz invece tratta gli urti sulla base della equazione di Boltzmann, a partire da un modello in cui gli atomi sono immaginati come sfere dure disposte a caso con densità n e gli elettroni si muovono urtando elasticamente contro gli atomi ma senza altre interazioni (questo modello di moto con urti è interessante di per se ed è noto come "modello di Lorentz" $\rightarrow$). In entrambe le teorie si suppone che gli elettroni abbiano una distribuzione di velocità data da una maxwelliana. Invece nella teoria di Sommerfeld si tiene conto della natura quantica degli elettroni e si suppone una distribuzione di Fermi Dirac sulle velocità, utilizzando poi le formule della teoria di Lorentz.

Sebbene i risultati delle teorie di Drude e Lorentz siano in buon accordo con la legge di Wiedeman e Franz, ossia prevedano il corretto rapporto (almeno come ordine di grandezza) fra le conducibilità termica ed elettrica, i risultati per i valori separati di σ_{el} e σ_{term} hanno dipendenza errata dalla temperatura T per $T \rightarrow 0$, e quindi il successo delle teorie è da considerarsi casuale. Diversamente avviene invece per la teoria di Sommerfeld, che fornisce anche la dipendenza qualitativamente corretta sia per σ_{el} che per σ_{term} per $T \rightarrow 0$.

E' fondamentale osservare che la teoria completa della conduttività elettrica e termica deve necessariamente contenere un' analisi delle interazioni fra elettroni e reticolo (e fra elettroni ed elettroni) perchè sono queste interazioni che danno luogo a un cammino libero medio $l < \infty$: i modelli in cui tali interazioni vengono trascurate danno luogo a sistemi con conducibilità elettrica e termica infinite, come segue facilmente dalla teoria dei gas perfetti classici o quantici.

Una importante conseguenza è che, se un modello prevede che i portatori di elettricità occupino tutti i possibili livelli di energia $\leq \varepsilon_0$ e che gli altri livelli possibili abbiano energia $\geq \varepsilon_0 + \Delta$ con $\Delta > 0$ allora le collisioni con gli atomi del reticolo potranno aver luogo solo se l' energia termica $\sim k_B T$ di questi ultimi è superiore a Δ. Ne segue chi in tali modelli, a dispetto della esistenza di interazioni fra portatori di elettricità e atomi del cristallo, tali interazioni non producono diffusione degli elettroni (cioè $l = \infty$) se $k_B T \ll \Delta$ e quindi questi modelli mostrano conducibilità elettriche infinite a temperature abbastanza basse dando luogo a fenomeni di

superconduttività, v. Modello BCS.

Bibliografia: Becker, R.: *Teoria della elettricità*, Sansoni, Firenze, 1950, §35,36,41.

Conducibilità, termica: in base alla legge di Fourier la quantità di calore che attravesa un elemento ideale di superficie $d\sigma$ tracciato in un continuo isotropo è proporzionale al prodotto scalare fra la normale $\underline{n}$ alla superficie e il gradiente della temperatura: $dQ = -\kappa \operatorname{grad} T \cdot \underline{n}\, d\sigma$. Il segno $-$ tien conto che il calore fluisce dalla faccia rivolta verso la regione a temperatura più alta all' altra. La costante κ è la conducibilità termica. Nei mezzi anisotropi la conducibilità è un tensore κ_{ij} e la legge di Fourier diviene: $dQ = -\sum_{i,j=1}^{3} \kappa_{ij} n_i (\partial T/\partial x_j)\, d\sigma$

Corpo nero: cavità con pareti a temperatura fissata, all' interno della quale si trova un campo di radiazione elettromagnetica in equilibrio termico con le pareti stesse, (v. Statistiche quantiche, MSC).

Correlazione, funzioni di C.: data una distribuzione di probabilità che descrive un sistema di particelle distribuite con posizioni e velocità casuali, le funzioni di C. sono funzioni $\rho(x_1, x_2, \ldots, x_n)$ di n punti nello spazio delle posizioni e velocità che danno la densità della probabilità di trovare una particella in ciascuno degli elementi di volume dx_1,...,dx_n attorno a $x_1, \ldots, x_n$. L' insieme delle funzioni di C. (al variare di n) fornisce una descrizione statistica completa dello stato del sistema. In termini delle funzioni di C. si possono esprimere le *distribuzioni locali*: se Λ è un volume finito le distribuzioni locali relative al volume Λ sono le densità $f_\Lambda(x_1, \ldots, x_n)/n!$ della probabilità per l' evento in cui nel volume Λ si trovano esattamente n particelle, e tali particelle sono esattamente negli elementi di volume dx_1,...,dx_n attorno a $x_1, \ldots, x_n$. La funzione di C. a due punti ($n = 2$) è molto studiata; nella MSC in termini di essa sono calcolabili, data la temperatura, l' energia interna e la lunghezza di C. (v.). Negli stati di equilibrio termodinamico la dipendenza dagli impulsi è banale; infatti, se si pone $x_i = (\underline{p}_i, \underline{r}_i)$, la funzione di C. fattorizza in una funzione $\rho(\underline{r}_1, \ldots, \underline{r}_n)$ o $f_\Lambda(\underline{r}_1, \ldots, \underline{r}_n)$ delle sole coordinate spaziali e in una funzione delle sole $\underline{p}_i$ e questo secondo fattore è una gaussiana, perchè la distribuzione degli impulsi è maxwelliana.

Correlazione, funzioni di C. spaziali: si ottengono dalle funzioni di C. integrando le cordinate di impulso (da cui, negli stati di equilibrio termodinamico dipendono banalmente), (v. funzioni di C.). La funzione di C. a due punti integrata sulle coordinate di impulso è una funzione $\rho(\underline{r}_1, \underline{r}_2)$ delle posizioni $\underline{r}_1, \underline{r}_2$ dei due punti e, negli stati di equilibrio termodinamico, descriventi fasi pure, tende, per $r = |\underline{r}_1 - \underline{r}_2| \to \infty$, al quadrato ρ^2 della densità (v. Fattorizzazione, Fase pura). La funzione $h(r) = \rho(\underline{r}_1, \underline{r}_2) - \rho^2$ si dice funzione di C. troncata a due punti. Nei sistemi a punto critico normale e con interazioni a corta portata questa funzione è, al punto critico,

una funzione che si comporta, per $r \to \infty$, come $r^{-(d-2+\eta)}$, se d è la dimensione dello spazio ed η è uno degli esponenti critici (v. Esponenti Critici, Fenomeni Critici); al di fuori del punto critico è, in questi sistemi, una funzione che decresce esponenzialmente.

Costante dei Gas: è il rapporto $R = PV/T$ in una grammomolecola di gas perfetto; si ha $R = N_A k_B$ ove N_A è il numero di Avogadro (v.) e k_B è la costante di Boltzmann (v.); $R = 8.31\, J\, ^c ircK^{-1}$. L' ordine delle definizioni è: prima si definisce la temperatura assoluta (dal secondo principio della termodinamica), (v.), poi la si identifica con la temperatura definita dal termometro a gas perfetto e si usa la legge di Avogadro per dedurre che PV/T è, in un gas perfetto, proporzionale al numero di molecole; e, infine, si definisce una grammomolecola di idrogeno come $2\, g$ di idrogeno e quindi $R = PV/T$ è ottenuto misurando P, V, T in un gas di $2g$ di idrogeno in uno stato di gas perfetto (ad esempio in condizioni normali). Il numero di Avogadro si deduce, infine, misurando la costante di Boltzmann (v.).

Cuore duro : se $2r$ è il raggio della sfera nella quale il potenziale di interazione di una molecola con una molecola identica vale $+\infty$, allora la sfera di raggio r attorno ad una molecola è il *cuore duro* fra le due molecole. La distanza $2r$ è quindi la distanza minima alla quale si possono avvicinare due molecole a cuore duro. In sistemi con più specie di molecole si possono definire le distanze $2r_{ij}$ di minimo avvicinamento fra le molecole della specie i a quelle della specie j. Le distanze r_{ij} si dicono raggi dei cuori duri fra le due specie. Una interazione si dice a *cuore duro* fra la specie i e la j se $r_{ij} > 0$. Se il potenziale di interazione fra specie i e specie j vale 0 a meno che la distanza sia inferiore a r_{ij} si dice che fra le due specie esiste una interazione di solo cuore duro.

Curie-Weiss, teoria di C.W.: v. teoria del Campo medio, Transizioni di fase.

Debye, teoria di D. del calore specifico dei cristalli: calcola il calore specifico di un cristallo in base al modello che lo schematizza come un cristallo perfettamente elastico in un volume di lato L nel quale si trova anche, nel caso di un cristallo conduttore, un gas di elettroni liberi (gli elettroni di conduzione). Il calore specifico C_V è la somma $C_V^e + C_V^c$ dei contributi dovuti agli elettroni e alle vibrazioni del cristallo. Il primo contributo è calcolato dalla teoria di un gas perfetto quantico di Fermi; il risultato è che il contributo in questione, C_V^e, al calore specifico è $(3R/2)$ per mole se $T \gg \Theta_e$ ove Θ_e è una temperatura assai alta, nei metalli ordinari, rispetto al punto di fusione del metallo (ad esempio è di $\sim 1.35\, 10^{5^\circ} K$ per il rame), mentre se $T \ll \Theta_e$ allora $C_V \simeq (3R/2)(\pi^2 T/3\Theta_e) \equiv \gamma T$. Questo fornisce la spiegazione del perchè a temperatura ambiente gli elettroni di un metallo non contibuiscono, proporzionalmente al numero dei loro gradi di libertà, al calore specifico come richiederebbe il principio di equipartizione dell' en-

ergia della MSC. Il secondo contributo, C_V^c, viene calcolato dalla teoria del gas perfetto di bosoni: infatti un cristallo perfettamente elastico è formalmente analogo ad un corpo nero e i quanti di vibrazione vengono chiamati *fononi*, v.; per semplificare il calcolo la teoria di Debye pone le due seguenti approssimazioni:
1) si immagina che i fononi non siano divisi in bande (acustiche e ottiche) ma che siano descritti da un impulso $\underline{k} = 2\pi L^{-1}\underline{n}$, ove $\underline{n}$ è un vettore a componenti intere. Si suppone che $|\underline{k}| \leq k_M$ e, per consistenza, si determina k_M in modo che il numero di tali modi (cioè di tali vettori $\underline{k}$) sia esattamente il numero di gradi di libertà del sistema di oscillatori (ossia $3N$ se N è il numero di atomi): questa è una approssimazione che sarebbe corretta solo nel caso di cristalli monoatomici con reticolo cubico semplice (caso in cui esiste solo la banda acustica). Si trova $k_M = (6\pi^2 n)^{1/3}$ dove n è la densità numerica.
2) si suppone che l' energia dei fononi di impulso $\underline{k}$ sia $E(\underline{k}) = \hbar|\underline{k}|v_0$, ove v_0 è la velocità del suono e $\hbar$ la costante di Planck divisa per 2π. Questa ipotesi è corretta solo per $|\underline{k}|$ molto piccolo (e per la sola banda acustica) in cristalli in cui si può supporre che la velocità del suono sia indipendente dalla direzione di propagazione.

Si deduce allora che se $T \ll \Theta' = \hbar v_0 k_M / k_B$, ove k_B è la costante di Boltzmann, il calore specifico per mole è $C_V^c = \alpha T^3$, con $\alpha = 2\pi^2 k_B^4 / 5\hbar^3 v_0^2$, mentre se $T \gg \Theta'$ si ha $C_V^c = 3R$, cioè la legge di Dulong e Petit, v. Il quadro completo che viene fornito dalla teoria di Debye è il seguente: si definiscono tre scale di temperatura; $\Theta_e, \Theta', \Theta"$, con $\Theta" = (\gamma/\alpha)^{1/2}$. Nei metalli ordinari si ha $\Theta_e \gg \Theta' \gg \Theta"$. Ad es. nel rame si trova che $\Theta' \sim 350°K$, $\Theta" \sim 5°K$ e se $T \ll \Theta"$ il valore del calore specifico è dominato dal contributo degli elettroni di conduzione; per $\Theta" \ll T \ll \Theta_e$ è invece dominato dalle vibrazioni reticolari e questa regione si divide nella regione in cui $\Theta" \ll T \ll \Theta'$ in cui $C_V \sim \alpha T^3$ e nella regione $\Theta' \ll T \ll \Theta_e$ in cui C_V è $3R$ con una correzione lineare; se ν è il numero di elettroni di conduzione per atomo si ha dunque:

$$\begin{aligned} C_V &= \nu\gamma T \qquad && T \ll \Theta" \\ C_V &= \alpha T^3 \qquad && \Theta" \ll T \ll \Theta' \\ C_V &= 3R + \nu\gamma T \qquad && \Theta' \ll T \ll \Theta_e \end{aligned} \tag{1}$$

e la regione $T \gg \Theta_e$ è nei metalli ordinari priva di significato perchè Θ è superiore alla temperatura di fusione (ove non ha più senso il modello). Però la costante ν va pensata come un parametro empirico, vista la crudezza del modello. La presenza degli elettroni di conduzione è pertanto rivelabile sperimentalmente a bassa e ad alta temperatura (mentre a temperature

intermedie è praticamente impossibile distinguerla dal contributo delle vibrazioni perchè quest' ultimo dovrebbe essere noto con un' approssimazione che la teoria può difficilmente raggiungere). Le approssimazioni caratteristiche della teoria di Debye sono eseguite solo per semplificare il calcolo che resterebbe, in linea di principio, eseguibile con la precisone desiderata: il comportamento descritto nella (1) resterebbe lo stesso con valori diversi per le costanti. Dunque la vera approssimazione di natura fisica della teoria è il modello del cristallo come perfetto e del gas di elettroni di conduzione come immerso nel cristallo senza alcuna interazione, nè fra elettroni ed elettroni, nè fra elettroni e cristallo. La teoria di Debye vera e propria teneva conto del solo contributo delle vibrazioni, e si applicava dunque propriamente ai cristalli non metallici (in cui $\nu = 0$ e $\Theta_e = +\infty$), mentre la teoria del contributo elettronico fu aggiunta da Sommerfeld.

Debye, teoria dello schermaggio di D.: in un gas rarefatto, ad alta temperatura, ed elettricamente neutro di particelle cariche, di carica $\pm q$. Le cariche del gas si dispongono attorno a una carica q, di prova, in modo che il potenziale elettrostatico generato da quest' ultima a distanza r appaia come dato da $q\, r^{-1} e^{-r/r_D}$, con $r_D > 0$, invece che $q\, r^{-1}$. La lunghezza r_D è la *lunghezza di Debye* che è data da $r_D = (\beta \rho q^2)^{-1/2}$, ove ρ è la densità numerica, $\beta = 1/k_B T$ con T la temperatura assoluta e k_B la costante di Boltzmann. La teoria si basa su una approssimazione di campo medio, secondo la quale la densità della nuvola di carica che si dispone attorno ad una carica q posta nel gas fa sì che, in ogni punto $\underline{r}$ del gas attorno alla carica, si osservi un potenziale elettrico $\varphi(\underline{r})$ incognito; e allora dalla MSC si può calcolare la densità di carica in $\underline{r}$ poichè la probabilità di trovare una carica $\pm q$ è proporzionale a $e^{-(\pm q)\beta\varphi(\underline{r})}$ (trascurando le fluttuazioni) e quindi tale densità è $q\rho \tanh(q\varphi(\underline{r}))$. Allora dall' equazione di Poisson (v.) dell' elettrostatica si vede che φ deve verificare, per consistenza, la $\Delta\varphi = -4\pi\rho q \tan(\beta q \varphi))$. La formula di Debye segue risolvendo questa equazione sotto la ulteriore approssimazione che sostituisce la tangente con il suo argomento. La approssimazione è accettabile se $r \gtrsim r_D$ e se la temperatura è abbastanza alta perchè sia $\beta q^2/r_D << 1$ (ossia se $q^3(\rho\beta^3)^{1/2} \ll 1$; quest' ultima relazione implica anche, come si può verificare, che il numero di cariche contenute nella sfera di raggio r_D è molto grande, che è un' altra necessaria proprietà di consistenza).

Decomposizione ergodica: lo spazio delle fasi di un sistema dinamico non ergodico può essere pensato come unione di sottoinsiemi (trasformati in se stessi dall' evoluzione temporale, cioè invarianti) su cui l' evoluzione agisce in modo ergodico. La possibilità di una tale decomposizione è evidente nel caso di un sistema dinamico con spazio degli stati finito (come avviene nei sistemi dinamici studiati numericamente) ma richiede un minimo di sofisticazione matematica per essere correttamente formulata nel

caso di sistemi dinamici con spazio delle fasi continuo.

Degenerazione, nella MSC: è la situazione in cui si viene a trovare un sistema meccanico classico quando se ne calcolano le proprietà termodinamiche, a mezzo dela MSC, dividendo lo spazio delle fasi in cellette di dimensioni h^{3N}, (h= costante di Planck) e si trovano, in corrispondenza dei parametri termodinamici considerati, risultati sensibilmente dipendenti dal valore di h (intorno a $h = 6.62\,10^{-27}\,erg\,sec$). In questi casi la MSC non è più affidabile e occorre utilizzare la MSQ per studiare le proprietà del sistema, (v. Condensazione di Bose, Sfera di Fermi, MSC, Statistiche quantiche).

Demonietto: vezzeggiativo di Demone, v. Maxwell.

Demone: programma di elaborazione, di solito attivato automaticamente all' accensione di un ordinatore elettronico e dedicato all' esecuzione di varie operazioni, spesso di controllo sull' uso dell' elaboratore stesso. Più in generale: essere in grado di eseguire operazioni "microscopichè, ripetitive e noiose, ma gravide di conseguenze. Appare spesso nel corso di simulazioni numeriche nella Meccanica Statistica. Da piccolo fu evocato da Maxwell, (v.),. Ne esistono anche forme degeneri e dannose, dette però più propriamente "virus".

Diffusione, processo di: moto aleatorio in cui il quadrato della distanza dal punto di partenza cresce, in media, proporzionalmente al tempo. Il coefficiente di proporzionalità D si chiama coefficiente di diffusione. Se le variazioni delle posizioni in intervalli di tempo disgiunti $\Delta t_1, \Delta t_2, \ldots$ sono variabili aleatorie indipendenti con distribuzione di probabilità gaussiana e con dispersioni $D\Delta t_1, D\Delta t_2, \ldots$ allora il processo di diffusione è *normale* e si dice che le traiettorie eseguono il *moto browniano*, (v.).

Dimensione di informazione: nozione di dimensione frattale, legata alla dimensione di correlazione. Se A è un insieme limitato in R^n e se μ è una distribuzione di probabilità definita su A, la dimensione di informazione di A rispetto alla distribuzione μ è l' estremo inferiore delle dimensioni di Hausdorff dei sottoinsiemi B di A che hanno probabilità 1 ($\mu(B) = 1$). Il teorema di L. Young afferma che se con μ–probabilità 1 i punti $x \in A$ sono tali che $\lim_{r\to 0} \frac{\log \mu(\text{sfera di centro } x \text{ e raggio } r)}{\log r} = \alpha$ allora la dimensione di informazione di A rispetto a μ è α. Se S è una trasformazione due volte differenziablile nell' intorno di A e $n = 2$ la dimensione di informazione δ è legata agli esponenti di Lyapunov e alla entropia $s(S)$ del sistema dinamico (S, μ) dalla formula di L. Young: $\delta = s(S)(|\lambda_1|^{-1} + |\lambda_2|^{-1})$, se gli esponenti di Lyapunov sono non nulli e di segno opposto.

Dimensione frattale: estensione ad insiemi limitati arbitrari $A \subset R^n$ della nozione di dimensione di una porzione di superficie regolare. La nozione non è unica: ne esistono varie definizioni, equivalenti solo nei casi più semplici. Le definizioni possono essere *statiche* o *dinamiche*. Nel primo caso dipendono esclusivamente dalla sola struttura dell' insieme A; nel sec-

ondo caso dipendono non solo da A ma anche dal modo in cui si pensa di generare l' insieme A stesso: e di solito le definizioni riguardano insiemi A generati da traiettorie di una "dinamica" S su R^n, ossia sono insiemi A consistenti nei punti $S^k x$, $k = 0, 1, \ldots$, con x punto in R^n prefissato e con S trasformazione di R^n in se, e nei loro punti di accumulazione. Rientrano in questa classe di insiemi gli attrattori strani dei sistemi dinamici, v. Caos. Strettamente parlando, dunque, la dimensione frattale dinamica è una funzione della coppia A, S. Esempio tipico di dimensione frattale è la *dimensione di Hausdorff*, v.; esempio tipico di dimensione frattale dinamica è la dimensione di correlazione, v.

Dimensione di Correlazione: nozione di dimensione frattale, v., dinamica. Se A è la chiusura di una traiettoria $x \to S^k x$, $k = 0, 1, \ldots$ generata dalle iterate di una trasformazione S regolare (differenziabile a tratti) di R^n in se, si considera un tratto della traiettoria di lunghezza N, ossia il tratto con $0 \leq k \leq N-1$. Si conta la frazione $\nu_N(j) = \frac{1}{N}\mathcal{N}_N(j, \delta)$ degli N punti del tratto di traiettoria che cadono in una sferetta di diametro δ e centro il j-mo punto della traiettoria, $S^j x$, e il suo valor medio su j: $\frac{1}{N}\sum_{j=0}^{N-1} \nu_N(j) = \frac{1}{N^2}\sum_{j=0}^{N-1} \mathcal{N}_N(j, \delta)$. Si considera poi il limite $\nu(\delta)$ per $N \to \infty$, se esiste. La dimensione di correlazione è α_c se $\frac{\nu(\delta)}{\delta^\alpha} \xrightarrow[\delta \to 0]{} 0$ per $\alpha < \alpha_c$ e $\frac{\nu(\delta)}{\delta^\alpha} \xrightarrow[\delta \to 0]{} \infty$ per $\alpha > \alpha_c$.

Dimensione di Hausdorff: nozione di dimensione frattale, v. Se $A \subset R^n$ è un insieme limitato se ne considerano tutti i ricoprimenti mediante un numero finito di insiemi chiusi di diametro $\leq \delta$, con δ prefissato. Se $\mathcal{C}_\delta$ è un tale ricoprimento si considera, per $\alpha > 0$, la quantità:

$$\mu_\alpha(A) = \lim_{\delta \to 0} \inf_{\mathcal{C}_\delta} \sum_{C \in \mathcal{C}_\delta} (\operatorname{diam} C)^\alpha$$

e si dimostra che esiste un valore α_c tale che $\mu_\alpha(A) = +\infty$ se $\alpha < \alpha_c$ e $\mu_\alpha(A) = 0$ se $\alpha > \alpha_c$. Il valore $\mu_{\alpha_c}(A)$ definisce la *misura di Hausdorff* e α_c definisce la *dimensione di Hausdorff* di A. La misura di Hausdorff può essere 0, ovvero un numero positivo, ovvero $+\infty$. Un esempio notevole è fornito dagli insiemi di Cantor A_m, $m \geq 3$: sono i punti di $[0, 1]$ che nella loro rappresentazione in base m non contengono la cifra 1. Si trova che la dimensione di Hausdorff di A_m è $\alpha_c = \frac{\log(m-1)}{\log m}$ e misura frattale 1. Invece l' insieme $\cup_{m \geq 3}^{\infty} A_m$ ha dimensione di Hausdorff 1 e misura di Hausdorff 0. Ovviamente l' intero intervallo $[0, 1]$ ha misura e dimensione 1. Un esempio meno banale di insieme frattale è un tratto di traiettoria browniana $t \to \omega(t)$, $t \in [0, 1]$, in R^n, v. moto browniano: con probabilità 1 questo insieme ha dimensione 2. Sono possibili altre definizioni simili di dimensione frattale ottenute ponendo vincoli sul tipo di ricoprimenti $\mathcal{C}_\delta$ da considerare nella definizione. Ad esempio la *dimensione a scatola* consiste nel considerare solo ricoprimenti consistenti in cubetti con diagonale δ e

sostituire il limite su $\delta \to 0$ con il limite superiore (perchè in generale il limite potrebbe non esistere, a differenza del caso di Hausdorff in cui si può mostrarne l' esistenza): con questa definizione si troverebbe che i numeri razionali hanno dimensione 1, mentre la loro dimensione di Hausdorff è nulla. Migliore risultato si otterrebbe vincolando i ricoprimenti a consistere di cubetti di diagonale $\leq \delta$ (invece che $= \delta$), o ad essere sferette di diametro $\leq \delta$, o ad essere parallelepipedi di base con diagonale δ e altezza $c\delta^2$, con $c > 0$ fissato. Sono tutte definizioni che sono state considerate in varie applicazioni.

Dimensione di Lyapunov: nozione di dimensione frattale (dinamica). Se (A, S, μ) è un sistema dinamico ergodico con $A \subset R^n$ limitato e S differenziabile a tratti, si considerano gli esponenti di lyapunov di (A, S, μ), $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ e si costruisce la funzione lineare a tratti che nei punti $\alpha = 0, 1, 2, \ldots$ vale $0, \lambda_1, \lambda_1 + \lambda_2, \lambda_1 + \lambda_2 + \lambda_3, \ldots$. La dimensione di Lyapunov è definita dal valore di α in cui questa funzione lineare a tratti si annulla; se la funzione non si annulla per $\alpha \leq n$ allora si pne $\alpha_c = n$. In generale la dimensione di Lyapunov non è inferiore alla dimensione di informazione di A rispetto a μ: si congettura che siano uguali.

Dinamica discretizzata: trasformazione di coordinate che, applicata ad un dato iniziale e iterata n volte, approssima la soluzione di una equazione differenziale dinamica valutata al tempo $n\tau$, con lo stesso dato iniziale. Il parametro τ è di solito liberamente scelto ed è detto *passo di integrazione*. Si richiede sempre che nel limite $\tau \to 0$ gli errori di approssimazione si annullino. Ad esempio l' equazione $\dot{\underline{x}} = \underline{f}(\underline{x})$, $\underline{x}(0) = \underline{x}_0$ può essere discretizzata dalla trasformazione $\underline{x}' = \underline{x} + \tau \underline{f}(\underline{x})$ e ponendo $\underline{x}_{n+1} = \underline{x}_n + \tau \underline{f}(\underline{x}_n)$, $\underline{x}_0 = \underline{x}(0)$ si ha che $\underline{x}_n$ approssima $\underline{x}(n\tau)$ e tende ad esso per $\tau \to 0$ e $n\tau = t$ fisso, (discretizzazione, o metodo, di Eulero al primo ordine). Altro esempio è il metodo di Eulero al secondo ordine: $\underline{x}_{n+1} = \underline{x}_n + \tau \underline{f}(\underline{x}_n) + (\tau^2/2)(\underline{f}(\underline{x}_n) \cdot \text{grad})\underline{f}(\underline{x}_n)$. La discretizzazione è di grande importanza quando si vuole affrontare lo studio numerico di una equazione differenziale e la scelta del metodo di discretizzazione da usare è il primo problema che si deve affrontare. Quando l' equazione che si vuole discretizzare ha particolari simmetrie si cerca, ove possibile, di rispettarle nella discretizzazione: di solito questo conduce ad algoritmi più efficienti (ossia ad algoritmi che forniscono migliori approssimazioni a parità di quantità di calcoli o tempo macchina). Ad esempio volendo discretizzare un sistema di equazioni hamiltoniane si cerca di rispettare la struttura canonica delle equazioni. Se la funzione hamiltoniana ha la forma $H = \sum_{i=1}^{l} p_i^2/2 + V(x_1, \ldots, x_l)$ un metodo molto usato è di definire, posto $\partial_i V = \partial V / \partial x_i$:

$$\begin{aligned} p_i^{n+1} &= p_i^n - \tau\, \partial_i V(\underline{x}^n) \\ x_i^{n+1} &= x_i^n + \tau\, p_i^{n+1} \end{aligned}$$

che ha il pregio di definire una trasformazione fra $(\underline{x}^n, \underline{p}^n)$ e $(\underline{x}^{n+1}, \underline{p}^{n+1})$ che è canonica (v.); laddove, se nella seconda equazione si scrivesse τp_i^n si troverebbe il metodo di Eulero al primo ordine (che *non* definisce una trasformazione canonica).

Dinamica simbolica: se S è una trasformazione di un insieme $\underline{\omega}$, detto spazio delle fasi, in se stesso e se $\underline{\omega}$ viene suddiviso in $P_1, \ldots, P_s$ ad ogni traiettoria $S^i x$, $i = 0, 1, 2 \ldots$ si può associare una successione $\underline{\sigma} = (\sigma_i)_{i=0,1,2,\ldots}$ ove, per ogni i, σ_i è l'indice tale che $S^i x \in P_{\sigma_i}$. In questo modo ad ogni punto dello spazio delle fasi $x \in \underline{\omega}$ è associata una successione di simboli $\underline{\sigma} = (\sigma_i)_{i=0,1,2,\ldots}$, detta *storia* di x sulla partizione $P_1, \ldots, P_s$ dello spazio delle fasi $\underline{\omega}$. L' azione di S sui punti di $\underline{\omega}$ può essere vista semplicemente come la traslazione delle intere successioni che ne rappresentano le storie; la storia di Sx si ottiene da quella di x, evidentemente, traslando la storia di x di una unità verso sinistra e cancellando il primo simbolo: $\underline{\sigma}(Sx)_i = \underline{\sigma}(x)_{i+1}$, per ogni i. Dunque nei casi in cui le storie dei punti determinano univocamente i punti che le generano la conoscenza del codice $x \longleftrightarrow \underline{\sigma}(x)$ determina completamente l' azione di S: si dice, in questi casi, che la S, vista come azione τ di traslazione sulle storie dei punti è una *dinamica simbolica* associata alla dinamica S e la partizione $P_1, \ldots, P_s$ si dice *generante*. E' chiaro che la dinamica simbolica può essere interessante per un' applicazione solo se gli insiemi $P_1, \ldots, P_s$ che la generano possono essere scelti in modo che l' insieme delle successioni $\underline{\sigma}$ che corrispondono a qualche x è descrivibile in modo semplice. Ad esempio in certi sistemi dinamici (rari, ma importanti) si possono trovare partizioni $\mathcal{P} = (P_1, \ldots, P_s)$ tali che l' insieme dei punti di $\underline{\omega}$ può essere posto in corrispondenza biunivoca con l' insieme delle storie su $\mathcal{P}$ e inoltre le storie possibili sono semplicemente *tutte*, ovvero sono tutte quelle per cui valgono certe relazioni prefissate fra simboli che possono apparire l' uno di seguito all' altro nella successione, (dinamiche simboliche markoviane), (v. Entropia e Informazione). Un esempio *tipico* è fornito da una trasformazione S dell' intervallo $\underline{\omega} = [0, 1]$, definita da $Sx = 10\,x \bmod 1$ e osservata sugli intervalli $P_k = [\frac{k}{10}, \frac{k+1}{10}$, $k = 0, \ldots, 9$. É il caso più semplice e la storia di x diviene semplicemente la successione di numeri che rappresentano x in base 10 (occorre però, perchè la corrispondenza punto storia sia biunivocamente definita, che da $[0, 1]$ si tolgano i punti che in base 10 hanno una periodicità 9, e il punto 1 stesso). Il caso in cui $Sx = 10\,x \bmod 1$ è il caso più semplice). Questo œ[-1zmette in luce il fatto che la dinamica simbolica è una estensione del ben noto procedimento per la rappresentazione dei numeri reali a mezzo di simboli (come gli interi da 0 a 9). La nozione di dinamica simbolica può essere estesa facilmente a sistemi dinamici in cui S è invertibile.

Distribuzione delle velocità, di Maxwell–Boltzmann: in MSC è la legge che afferma che le componenti della velocità (o impulso) delle particelle di un sistema classico hanno, in equilibrio termodinamico, distribuzioni indipendenti e gaussiane. E la dispersione di ciascuna componente della velocità è data in termini della costante k_B di Boltzmann, della massa m delle particelle e della temperatura assoluta T da k_BT/m. Ossia la probabilità che una particella di massa m abbia una componente della velocità fra v e $v + dv$ è proporzionale a: $[\exp -(mv^2/2k_BT)]dv$ (e la costante di proporzionalità è $(2\pi k_BT/m)^{-1/2}$, in modo che la probabilità che la velocità abbia un qualsiasi valore sia 1). Questa legge non è valida nella Meccanica Statistica quantistica (v. MSC, Equipartizione e Critica).

Distribuzione canonica: elemento dell' insieme statistico per la descrizione di un sistema in equilbrio termodinamico in contatto con una riserva di calore a temperatura T con la quale il sistema non può scambiare altro che calore. Nel caso di un sistema omogeneo di N particelle che occupa il volume Λ la probabilità di una configurazione con energia E è allora proporzionale a $\exp(-\beta E)$ (detto peso statistico (v.)), ove $\beta = 1/k_BT$ e la densità $\rho = N/V$ sono due parametri che caratterizzano l' elemento dell' insieme canonico considerato, (v. Insiemi Statistici). Il logaritmo della somma Q dei pesi statistici di tutte le configurazioni, cioè della somma di partizione, ha il significato fisico di prodotto fra $-\beta = -1/k_BT$ e l' energia libera $F = U - TS$.

Distribuzione gran canonica: elemento dell' insieme statistico per la descrizione di un sistema in equilbrio termodinamico con una riserva di calore a temperatura T con la quale il sistema può anche scambiare particelle identiche a quelle costituenti il sistema. Nel caso di un sistema omogeneo che occupa il volume Λ la probabilità di una configurazione con N particelle ed energia E è allora proporzionale a $\exp(\beta\mu N - \beta E)$ (detto peso statistico (v.)), ove $\beta = 1/k_BT$ e μ (potenziale chimico, v.) sono due parametri che caratterizzano l' elemento dell' insieme gran canonico considerato, (v. Insiemi Statistici). Il logaritmo della somma Z dei pesi statistici di tutte le configurazioni, cioè la somma di partizione, ha il significato fisico di prodotto βPV fra $\beta = 1/k_BT$, pressione e volume.

Distribuzione microcanonica, elemento dell' insieme statistico per la descrizione di un sistema in equilibrio termodinamico ma isolato meccanicamente e termicamente, con energia compresa fra U e $U - DU$, ove DU è una grandezza macroscopica (di solito piccola) prefissata. Nel caso di un sistema omogeneo che occupa il volume Λ la probabilità di una configurazione con N particelle ed energia E fra $U - DU$ ed U, con U assegnato e DU prefissato, è allora costante, ossia ogni configurazione ha peso statistico 1, v. IS, Peso Statistico). Il logaritmo della somma $\mathcal{N}$ dei pesi statistici di tutte le configurazioni, cioè della somma di partizione, ha il significato

fisico di rapporto fra entropia S del sistema e costante di Boltzmann k_B.

Distribuzioni locali, di probabilità: (v. Correlazione).

Dulong-Petit, legge di: conseguenza della legge di equidistribuzione dell' energia (v. MSC). In base ad essa un solido cristallino semplice (ad esempio un metallo), dovrebbe avere calore specifico $C_V = 3nR$, ove n è il numero di moli e R la costante dei gas. E' ben verificata solo a alta temperatura, (v. MSC, Debye).

Equazione di van der Vaals: equazione di stato approssimata per un gas non perfetto, v. serie del Viriale, e anche v. Teoria di Campo medio, Esponenti Critici, Potenziale di Kac)

Equazioni algebriche risolubili: le equazioni algebriche risolubili in termini di radicali sono le equazioni di grado inferiore al quinto. Le equazioni di primo grado hanno la forma $x + b = 0$ e quindi sono risolubili. Quelle di secondo grado hanno la forma $x^2 + 2bx + c = 0$ e quindi hanno le due soluzioni $x = -b \pm \sqrt{b^2 - c}$. Le equazioni di terzo grado si riducono facilmente alla forma $x^3 + 3px - 2q = 0$, (a partire dalla forma generale $y^3 + ay^2 + by + c = 0$ si ponga $x = y - \frac{a}{3}$). L' equazione nella forma ridotta è risolta dalla formula di Cardano:

$$x_0 = (q + \sqrt{p^3 + q^2})^{1/3} + (q - \sqrt{p^3 + q^2})^{1/3}$$

e le altre due soluzioni si ottengono risolvendo l' equazione di secondo grado $P(x) = 0$ ove $P(x) = (x^3 + 3px + 2q)/(x - x_0)$ si calcola esplicitamente per divisione. Le equazioni di quarto grado si riducono alla forma $x^4 + 2ax^2 + 2bx + c = 0$ con una sostituzione lineare analoga a quella delle equazioni di terzo grado. Si osserva poi che l'equazione ridotta può essere scritta, qualunque sia il parametro ausiliario t, come:

$$(x^2 + a + t)^2 = 2tx^2 - 2bx + (t^2 + 2at - c + a^2)$$

e si può scegliere t in modo che il secondo membro sia un quadrato perfetto, ossia in modo che $b^2 - 2t(t^2 + 2at - c + a^2) = 0$, determinando t come soluzione di questa equazione di terzo grado (usando ad esempio le formule di Cardano). Se t_0 è una soluzione, l' equazione originale assume la forma $(x^2 + a + t_0)^2 = 2t_0(x - \frac{b}{2t_0})^2$ che da luogo a due equazioni quadratiche per x^2 e, quindi, alla soluzione completa della equazione, tramite formule che potrebbero essere scritte esplicitamente: le formule di Ferrari. La risolubiltà delle equazioni di terzo grado fu trovata da Tartaglia, ma pubblicata contro la sua volontà da Cardano. La risolubilità delle equazioni di quarto grado è dovuta a Ferrari, allievo di Cardano. La non risolubilità, in generale, delle equazioni di grado superiore al quarto a mezzo di radicali fu dimostrata da E. Galois.

Bibliografia: V. Gindinkin, ...

Equazioni DLR, equazioni verificate dalle correlazioni locali (v.) di uno stato di equilibrio termodinamico di Gibbs (v. Limite Termodinamico, Gibbs). Le iniziali stanno per R. Dobrushin, O. Lanford, D. Ruelle.

Equazioni di KS, equazioni verificate dalle funzioni di correlazione (v.) di uno stato in equilibrio termodinamico di Gibbs, (v. limite termodinamico). La loro soluzione per serie conduce agli sviluppi in serie (in serie della della densità o dell' attività) della equazione di stato, validi a bassa densità, (v. Gerarchia di BBGKY). Le iniziali stanno per J. Kirkwood, Z. Salsburg.

Equipartizione dell' energia, teorema: nella MSC e in uno stato di equilibrio termodinamico, ogni grado di libertà traslatorio contibuisce la quantità $k_BT/2$ all' energia media e ogni grado di libertà oscillatorio contribuisce k_BT, (v. MSC). Questa proprietà non è più vera nella MSQ e ne sono esempi illustri la teoria del corpo nero e la teoria dei calori specifici a bassa temperatura, v. MSC, Equipartizione e Critica, Statistiche Quantiche, Debye.

Ergodicità, proprietà di un sistema dinamico metrico $(\underline{\omega}, S, \mu)$: se E è un insieme invariante μ–misurabile, allora la misura di E è 0 ovvero 1, per l' etimologia v. Monodo.

Ergodo, termine usato da Boltzmann per denotare quello che oggi è l' insieme statistico microcanonico, (v. Distribuzione microcanonica, IS, MSC), per l' etimologia v. Monodo.

Esperimento numerico: esperimento i cui risultati consistono nei dati di uscita (numerici o grafici) di un programma di calcolo a mezzo di un elaboratore elettronico digitale o analogico, quando il programma è ideato in modo da risolvere (di solito solo approssimativamente) le equazioni del moto che modellano un sistema reale (o anche ideale). Da un punto di vista filosofico l' esperimento numerico è un esperimento che studia il sistema costituito dal programa stesso e la difficoltà teorica e pratica sta nella interpretazione della rilevanza dei risultati per lo studio di sistemi reali. La situazione non è dunque molto diversa da quella in cui ci si trova quando si adotta un modello matematico per la descrizione di un fenomeno fisico e lo si riesce a studiare in qualche dettaglio, (v. Termalizzazione).

Esponenti critici: numeri puri caratterizzanti la natura della singolarità della dipendenza di varie grandezze termodinamiche in funzione di altre nelle vicinanze del punto critico di una transizione di fase, ovvero descriventi (sempre nelle vicinanze del punto critico) la dipendenza delle funzioni di correlazione spaziali (v.) in termini della distanza fra i loro argomenti. Esempi tipici di esponenti critici si ottengono considerando un sistema omogeneo che presenta una transizione liquido gas. Se T_c, P_c, ρ_c denotano la temperatura, la pressione e la densità del sistema al punto critico e se ρ_l, ρ_g denotano le densità del liquido e del gas coesistenti ad una data temperatura T, l' equazione di stato assume, nelle vicinanze del

punto critico e per $T = T_c$, la forma $P - P_c \propto \pm|\rho - \rho_c|^\delta$. Ovvero, sempre nelle vicinanze del punto critico ma per $\rho = \rho_c$, $T > T_c$, è tale che $(\partial p/\partial\rho)_{\rho=\rho_c} \propto (T - T_c)^\gamma$ e δ, γ sono due *esponenti critici*. Similmente il calore specifico a volume costante C_V assume, vicino al punto critico e per $\rho = \rho_c$, la forma $C_V \propto |T - T_c|^\alpha$ e $\rho_l - \rho_g \propto (T_c - T)^\beta$, ove α, β sono altri esponenti critici. Un altro importante esponente critico, μ, è legato alla tensione superficiale σ fra le due fasi coesistenti: questa si annulla al punto critico come $\sigma \propto (T_c - T)^\mu$. Esempi del secondo tipo di esponenti critici sono gli esponenti η e ν legati al comportamento per $r \to \infty$ della funzione di correlazione spaziale a due punti (v.) $h(r) = \rho(\underline{r}_1 - \underline{r}_2) - \rho^2$, ove $r = |\underline{r}_1 - \underline{r}_2|$: se $T \neq T_c$ e se d denota la dimensione dello spazio (di solito $d = 3$) si ha $h(r) \propto r^{-(d-2)} e^{-\kappa r}$ (v. Ornstein–Zernike) e $\kappa \propto |T_c - T|^\nu$, mentre se $T = T_c$ si ha $h(r) \propto r^{-(d-2+\eta)}$. Gli esponenti critici non sono sempre indipendenti e le relazioni che li legano si dicono *legge di scala*, (v.). Il valore numerico degli esponenti critici è molto sensibile alla dimensione dello spazio e alla portata del potenziale. Nelle teorie di campo medio, ossia in sistemi con interazione a lunghissima portata i valori sono indipendenti dalla dimensione; la seguente tabella riporta i valori noti o congetturati "esatti" di alcuni esponenti critici nella teoria di campo medio e nel modello di gas su reticolo di Ising a 2 dimensioni (v.) e i valori sperimentali per la (usuale) transizione liquido gas con interazione a corta portata a $d = 3$:

campo medio	corta port., $d = 3$	Ising, $d = 2$
$\alpha = 0$ discontinuità	$\alpha = 0.11 \pm 0.01$	$\alpha = 0$ sing. logaritm.
$\beta = 1/2$	$\beta = 0.325 \pm 0.005$	$\beta = 1/8$
$\gamma = 1$	$\gamma = 1.235 \pm 0.005$	$\gamma = 7/4$
$\delta = 3$	$\mu = 1.28 \pm 0.05$	$\delta = 15$
$\mu = 3/2$		$\mu = 1$
		$\eta = 1/4$

(v. Punto critico, Fenomeni critici, Leggi di scala, Transizioni di fase). Bibliografia: Rowlinson, J.: prefazione alla edizione critica della tesi di van der Waals, J.: *On the continuity of the gaseous and liquid states*, Studies in Statistical Mechanics, vol. XIV, J. Lebowitz series editor, North Holland, New York, 1988.

Esponenti di Lyapunov (locali): numeri che misurano l'azione di espansione e contrazione dei segmenti infinitesimi sotto l'azione delle ite–

rate di una trasformazione S, regolare (differenziabile a tratti e localmente invertibile) di R^n in se, definita nell' intorno di un insieme A chiuso limitato e S–invariante (ossia $S\,A \subset A$). L' azione di espansione di S^k, $k = 0, 1, \ldots$, nel punto $y \in A$ è descritta a mezzo della matrice M_k che dà la trasformazione lineare che trasforma un segmento infinitesimo $d\underline{l}$ uscente da y nella sua immagine $S^k d\underline{l} = M_k d\underline{l}$, che è pure un segmento infinitesimo uscente da $S^k y$. Se $d\underline{l}_1, \ldots, d\underline{l}_p$ sono p segmenti infinitesimi linearmente indipendenti uscenti da y, si considera il parallelepipedo da essi generato attorno a y e quello generato dalle loro immagini $M_k d\underline{l}_1, \ldots, M_k d\underline{l}_p$ attorno a $S^k y$. Sia $f_{p,k}(y)$ il massimo del rapporto fra i volumi di questi parallelepipedi (al variare dei paprallelepipedi stessi) e si ponga $\Lambda_p(y) = \lim_{k\to\infty} \frac{1}{k} \log f_{p,k}(y)$, se il limite esiste. Si pone, ricorsivamente, $r_1(y) = \Lambda_1(y)$ e, per $j \geq 2$, $r_j(y) = \Lambda_j(y)\,\Lambda_{j-1}^{-1}(y)$. Si ha $r_1(y) \geq r_2(y) \geq \ldots$. Se $\hat{r}_1(y) > \hat{r}_2(y) > \ldots > \hat{r}_s(y)$ sono i valori distinti assunti dai numeri $r_j(y)$ e $\hat{r}_j(y)$ appare ripetuto $n_j \geq 1$ volte, si dice che i coefficienti di Lyapunov di S in y sono $r_1(y), r_2(y), \ldots$ ovvero sono $\hat{r}_1(y), \hat{r}_2(y), \ldots$ con molteplicità rispettive $n_1, n_2, \ldots$. Gli esponenti di Lyapunov sono i logaritmi dei coefficienti di Lyapunov $\lambda_j = \log r_j, \hat{\lambda}_j = \log \hat{r}_j$. L' insieme dei punti $y \in A$ nei quali sono definiti gli esponenti di Lyapunov sarà denotato $L_0(A)$. Questo insieme non solo è, molto in generale, non vuoto ma ha probabilità 1 *rispetto a qualunque distribuzione di probabilità S–invariante* μ (ossia tale che $\mu(E) = \mu(S^{-1}E)$ per ogni insieme chiuso E) che dia probabilità nulla all' insieme dei punti in cui S non è differenziabile e ai punti che evolvendo sotto azione di S si avvicinano "troppo rapidamente" alle singolarità di S (condizioni banali se S è differenziabile, ma tecnicamente un pò complesse nel caso generale): è il teorema di S. Pesin. *Non* si deve credere che gli esponenti di Lyapunov siano indipendenti da $y \in L_0(A)$. Sono però costanti del moto ossia $\lambda_j(y) \equiv \lambda_j(Sy)$, e quindi se μ è una misura invariante S–ergodica, v., e che dia probabilità nulla all' insieme dei punti in cui S non è differenziabile, allora scegliendo a caso un punto y con distribuzione μ si trovano esponenti locali di Lyapunov che *non* dipendono dal punto trovato. Ovviamente, però, cambiando μ fra le misure ergodiche si troveranno (in generale) valori diversi degli esponenti di Lyapunov. Gli esponenti di Lyapunov sono intuitivamente (proporzionali, con fattore infinito, agli) autovalori delle matrici $(M_k^* M_k)^{1/2}$ per $k = \infty$, e ci si può domandare se sia possibile definire anche una nozione corrispondente di autovettori. Sarebbe naturale pensare che ad ogni punto $y \in A$ si possa associare, almeno se gli esponenti di Lyapunov hanno molteplicità 1, una base di vettori $v_1(y), \ldots, v_n(y)$ tali che $M_k v_j(y) \propto e^{k\lambda_j(y)} v_j(S^k y)$. In generale però questo non si può dire, neppure se gli esponenti di Lyapunov sono due a due distinti, $s = n$. Tuttavia qualcosa di molto simile è in generale vero: consideriamo il sottoinsieme $L(A)$ di $L_0(A)$ dei punti y

per i quali è possibile definire s sottospazi lineari $V_1(y) \equiv R^n \supset V_2(y) \supset V_3(y) \supset \ldots \supset V_s(y)$ tali che $M_k V_j(y) = V_j(S^k y)$ e di dimensioni rispettive $n, n-n_1, n-n_1-n_2, \ldots$; inoltre se $d\underline{l}$ è un segmento infinitesimo in $V_j(y)/V_{j+1}(y)$ si ha $\lim_{k\to\infty} \frac{1}{k} \log \frac{|M_k d\underline{l}|}{|d\underline{l}|} = \hat{\lambda}_j(y)$, $j = 1, \ldots, s$. E' possibile mostrare che pure $L(A)$ ha probabilità 1 rispetto a qualsiasi distribuzione di probabilità invariante definita su A e che dia probabilità nulla ai punti in cui S non è differenziabile e a quelli che evolvono avvicinandosi "troppo rapidamente" alle singolarità di S (una condizione sufficiente è che $\int \mu(dy)||M_1(y)|| < \infty$, se $||M_1||$ è la norma della matrice M_1). Infine si può dare una nozione di esponenti di Lyapunov anche per sistemi dinamici continui $t \to S_t$, ove S_t è una trasformazione differenziablie a tratti definita nell' intorno di un insieme invariante A chiuso e limitato. A tale scopo si fissa una unità di tempo t_0 e si trova che gli esponenti di Lyapunov di S_{t_0} hanno la forma $\lambda_j(y) = t_0 \tilde{\lambda}_j(y)$ e quindi è naturale chiamare $\tilde{\lambda}_j(y)$ gli esponenti di Lyapunov del sistema dinamico continuo S_t. Si noti che questi esponenti non sono numeri puri ma hanno la dimensione di un inverso di un tempo. Per esempi si veda la v. Multifrattalità.

Fase Pura: stato di equilibrio termodinamico in cui le funzioni di correlazione spaziale (v. Correlazione Spaziale) godono della proprietà di fattorizzazione: $\rho(\underline{r}_1, \ldots, \underline{r}_n, \underline{r}_1{}' + \underline{R}, \ldots, \underline{r}_m{}' + \underline{R})$ tende per $\underline{R} \to \infty$ al prodotto $\rho(\underline{r}_1, \ldots, \underline{r}_n) \cdot \rho(\underline{r}_1{}', \ldots, \underline{r}_m{}'$ nel senso che la differenza fra le due funzioni tende a zero. Nei sistemi con interazione a corta portata l' avvicinamento a zero avviene di solito esponenzialmente su una scala di lunghezza (lunghezza di correlazione) κ^{-1} e si dice che il sistema non è al punto critico; se invece avviene secondo un andamento a potenza in $|\underline{R}|$ (cioè $\kappa^{-1} = \infty$) allora si dice che il sistema è al punto critico (v.).

Fase Mista: stato di equilibrio termodinamico in cui le funzioni di correlazione spaziale (v. Correlazione Spaziale) non godono della proprietà di fattorizzazione (v. Fase Pura).

Fasi, spazio delle: in Meccanica Classica indica lo spazio degli stati quando i suoi punti sono descritti a mezzo di coordinate canoniche, e quindi le equazioni del moto hanno forma hamiltoniana (v.). Per estensione si chiama spesso spazio delle fasi lo spazio degli stati di un sistema dinamico generico, anche non hamiltoniano.

Fattore di Boltzmann: peso statistico di una configurazione in un elemento di un insieme statistico. Ad esempio se si considera un elemento dell' insieme canonico, corrispondente alla temperatura T e a una interazione con hamiltoniana E, il F.B. della configurazione C è $e^{-\beta E(C)}$. Il rapporto fra i fattori di Boltzmann fornisce il rapporto fra le probabilità di due configurazioni in un data distribuzione di un insieme statistico (v. Peso statistico, MSC, IS).

Fattorizzazione: proprietà delle funzioni di correlazione spaziali nelle

fasi pure (v. Fase pura).

Feigenbaum, scenario di F. per lo sviluppodella turbolenza: v. Caos.

Fermi, sfera di F., energia di F.: le particelle di un gas perfetto di particelle con spin 1/2, verificanti la statistica di Fermi–Dirac, allo zero assoluto ed in equlibrio, si dispongono in modo da occupare tutti i livelli di energia più bassa con molteplicità 2. I livelli del gas perfetto dipendono solo dal momento $\underline{k}$ delle particelle che, se il sistema occupa un volume di lato L, ha componenti quantizzate come $k_i = n\hbar 2\pi/L$, (v. MSC, Statistiche Quantiche). Si vede dunque che se il sistema contiene N particelle di massa m allora vengono riempiti tutti i livelli corrispondenti agli impulsi contenuti nella sfera (di Fermi) di raggio p_F tale che $N = 2\frac{4\pi}{3}\frac{p_F^3}{(2\pi\hbar)^3}L^3$: p_F si dice momento di Fermi, e $p_F^2/2m$ è l' energia di Fermi. (v. Statistiche quantiche, MSC). Una definizione alternativa è basata sulla proprietà, equivalente, della trasformata di Fourier della matrice di densità ridotta a una particella (v.): tale trasformata a impulso $\underline{k}$ vale 1 se $|\underline{k}| < p_F$ e zero altrimenti; e il momento di Fermi può essere definito come il luogo dei punti in cui questa funzione ha una discontinuità. In un gas in cui le particelle interagiscono con un' interazione invariante per rotazione la definizione della sfera di Fermi è data in questo secondo modo (ma la funzione in questione non passa più da 1 a 0 bensi' ha una discontinuità che dipende dalla intensità dell' interazione): non ha infatti più senso parlare di livelli energetici delle singole particelle. La sfera di Fermi non esiste sempre nei sistemi in interazione, anche se piccola; si ritiene che la sfera esista se l' interazione fra le particelle è repulsiva e che non esista (come discontinuità in senso stretto) nel caso di interazioni attrattive, (v. MSC, Statistiche quantiche, superficie di Fermi).

Fermi, superficie di F.: nel caso di sistemi con interazione non invariante per rotazione (ad esempio elettroni in un cristallo) la sfera di Fermi diviene una superficie chiusa che può avere struttura topologica anche molto complessa, (v. Fermi, sfera di).

Fermione, nome che designa una particella che insieme ad altre particelle ad essa identiche verifica la statistica di Fermi–Dirac, (v.).

Fermi–Dirac, statistica di F.D.: è verificata da un sistema di N particelle identiche se gli stati quantici in cui tali particelle possono trovarsi sono descritti da funzioni d' onda antisimmetriche nelle coordinate delle particelle.

Fermi–Pasta–Ulam, esperimento: uno dei primi esperimenti numerici, il suo risultato fu la prima verifica che l' ipotesi di equipartizione dell' energia è in generale falsa anche in sistemi di interesse per la Fisica, (v. Termalizzazione, Esperimenti numerici).

Fourier, legge per la conduzione del calore: v. Conducibilità Termica.

Frazioni continue: rappresentazione dei numeri positivi a mezzo di suc-

cessioni di interi. Se $a > 0$ allora a può essere scritto in uno ed un solo modo come:

$$a = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \dots}} \tag{1}$$

ove, se $[x]$ = parte intera di x, la successione a_i si costruisce con il seguente algoritmo:

$$a \to \begin{cases} a_0 = [a] \\ \tilde{a}_1 = (a - a_0)^{-1} \to \begin{cases} a_1 = [\tilde{a}_1] \\ \tilde{a}_2 = (\tilde{a}_1)^{-1} \to \begin{cases} a_2 = [\tilde{a}_2] \\ \tilde{a}_3 = (\tilde{a}_2 - a_2)^{-1} \dots \end{cases} \end{cases} \end{cases} \tag{2}$$

Il numero a è razionale se e solo se per un certo j si ha $a_j = \infty$, ossia la frazione continua si tronca. La frazione continua si denota di solito con il simbolo $[a_0; a_1; a_2; \dots]$.

Le proprietà fondamentali delle frazioni continue dei numeri irrazionali sono formulate in termini delle successioni p_j, q_j definite in modo che p_j/q_j sia la frazione irriducibile che da il valore della frazione continua troncata al j-mo denominatore. Posto $p_0 = a_0, q_0 = 1, p_{-1} = 1, q_{-1} = 0$ le frazioni p_j/q_j, dette *convergenti* della frazione continua si costruiscono ricorsivamente:

$$\begin{aligned} p_j &= a_j p_{j-1} + p_{j-2} \qquad j \geq 1 \\ q_j &= a_j q_{j-1} + q_{j-2} \qquad j \geq 1 \end{aligned} \tag{3}$$

Valgono le seguenti proprietà:

1) $(q_j q_{j+1})^{-1} < |a - p_j/q_j| < q_j^{-2}$ e quindi $\lim_{j\to\infty} p_j/q_j = a$. Inoltre i convergenti con j pari sono approssimanti di a per difetto, e sono crescenti con j, e quelli con j pari lo sono per eccesso, e sono decrescenti.

2) se $q' < q$ e p/q è un convergente della fraziona continua di a allora $|q'a - p'| > |qa - p|$ per ogni p' e, viceversa, se una coppia p, q verifica questa proprietà allora p/q è un convergente di a per qualche j (*i convergenti sono i migliori approssimanti* razionali di un numero a).

3) se $\varepsilon(T)$=massimo intervallo fra coppie adiacenti di numeri dell' intervallo $[0, 1]$ della successione $ka - [ka]$, $k = 0, 1, \dots T$ e se si pone $\varepsilon_n = |aq_n - p_n|$ si ha:

$$\begin{aligned} q_n \leq T < q_n + q_{n-1} \qquad &\to \qquad \varepsilon(T) = \varepsilon_{n-1} \\ q_n + q_{n-1} \leq T < 2q_n + q_{n-1} \qquad &\to \qquad \varepsilon(T) = \varepsilon_{n-1} - \varepsilon_n \\ \dots \\ (a_{n+1} - 1)q_n \leq T < a_{n+1}q_n + q_{n-1} \ &\to \ \varepsilon(T) = \varepsilon_{n-1} - (a_{n+1} - 1)\varepsilon_n \end{aligned} \tag{4}$$

che fa vedere come si distribuiscono le *migliori approssimazioni* rispetto alle altre possibili (si ricordi che $a_{n+1}q_n + q_{n-1} = q_{n+1}$).

4) ogni irrazionale a soluzione di una equazione di secondo grado a coefficienti interi (*irrazionale quadratico*) ha una frazione continua le cui cifre divengono definitivamente periodiche: e viceversa se un numero ha la frazione continua definitivamente periodica è un irrazionale quadratico.
5) un numero irrazionale si dice *diofantino di tipo* α se:

$$\sup_{q,p\,interi} |qa - p|^{-1} q^{-\alpha} = C(\alpha) < \infty$$

e si mostra che i numeri che non sono diofantini di tipo $\alpha > 2$ formano un insieme di misura nulla sulla retta. Mentre quelli che sono diofantini di tipo $\alpha = 2$ formano un insieme di misura nulla che contiene però tutti gli irrazionali a le cui cifre della associata frazione continua sono uniformemente limitate da un qualche numero $M(a)$. Non esistono irrazionali di tipo $\alpha < 2$. In questo senso il numero *più irrazionale* in $[0, 1]$ è il numero la cui frazione continua è interamente costituita da $a_i = 1$, per $i > 0$: questo numero è la *sezione aurea* $a = (\sqrt{5} - 1)/2$: ha $\alpha = 2$ e $C(\alpha)$ è il minimo valore possibile per $C(2)$.
6) non è *nota* la frazione continua di alcun numero algebrico irrazionale che non sia un irrazionale quadratico (ad esempio $2^{1/3}$), nel senso che non si sa se le cifre delle frazioni di tali numeri hanno proprietà di regolarità tali da poterle descrivere per intero in un numero finito di passi. Alcune specialissime frazioni continue sono note, soprattutto grazie alle rappresentazioni di alcune funzioni elemntari in termini di frazioni continue con cifre non intere, come ad esempio la tangente iperbolica che si può scrivere come $\tanh x = [x; 3x; 5x; \ldots]$ e quindi $\tanh(1)$ ha la frazione continua $[1; 3; 5; 7; \ldots]$.
7) le frazioni continue giocano un ruolo molto importante nella teoria dei moti quasi periodici in quanto, per i risultati 2), 3), permettono di studiare come i punti della forma $ka - [ka]$ riempiono densamente l' intervallo $[0, 1]$

Frequenze , indipendenti in un moto quasi periodico: v. Moti quasi periodici.

Frequenza di una stringa in una successione di simboli: data una successione $\underline{\sigma} = (\sigma_i)_{i=0,1,\ldots}$ e data una stringa finita di n simboli si osserva il tratto di $\underline{\sigma}$ con indici fra 0 e N e si conta il numero di volte che j è tale che il tratto di successione $\sigma_j, \sigma_{j+1}, \ldots, \sigma_{j+n-1}$ coincide con la stringa data. Se questo numero, denotato M_N è tale che esiste il limite $\nu = \lim_{N\to\infty} M_N/N$ si dice che la stringa appare con ferquenza ν definita, nella successione $\underline{\sigma}$. Una successione è "a frequenze definite" se tutte le possibili stringhe finite hanno frequenza definita, v. Entropia e complessità.

Funzione di autocorrelazione, se $A(t)$ è il valore di una grandezza osservabile al tempo t la quantità $C(t) = \lim_{T\to\infty} T^{-1} \int_0^T A(\tau)A(t + \tau)d\tau$ è la funzione di autocorrelazione della grandezza A. Le funzioni di autocorrelazione sono legate ai coefficienti di trasporto. Ad esempio se $A(t)$ è una

componente della velocità di una particella di un gas omogeneo in equilibrio termodinamico la funzione di autocorrelazione consente di calcolare il coefficiente di (auto)diffusione+ come $D = 3\int_0^\infty C(\tau)d\tau$, v. Coefficienti di Trasporto, Diffusione, Autodiffusione).

Funzione di partizione: dato un elemento di un insieme statistico, o più in generale, data una distribuzione di probabilità, è la somma dei pesi statistici (v.) delle varie configurazioni o, più in generale, dei vari eventi. Come tale dipende dalla convenzione adottata per la definizione dei pesi statistici (che sono definiti a meno di un fattore comune che può dipendere da qualsiasi quantità, imn particolare dai parametri termodinamici che individuano gli elementi dell' insieme statistico). Le convenzioni usuali adottate nel caso degli insiemi statistici della Meccanica Statistica sono tali che il logaritmo della funzione di partizione ha in quei casi un semplice significato fisico, (v. Distribuzione canonica, gran canonica, microcanonica, IS).

Funzione zeta (di un sistema dinamico): è una funzione generatrice per il conteggio delle molteplicità N_m, $m = 1, 2, \ldots$ delle orbite periodiche di periodo m di un sistema dinamico (A, S) con $A \subset R^n$ chiuso e limitato e S trasformazione continua di A in se. Si definisce, per s complesso: $\zeta(s) = \exp\sum_{m\geq 1} \frac{1}{m} N_m e^{-sm}$; e se ω è una generica orbita periodica di periodo $m(\omega)$ si ha: $\zeta(s) = \exp\sum_\omega e^{-sm(\omega)}$. Se P è l' insieme delle "orbite periodiche prime", ossia delle orbite periodiche che nel periodo $m(\omega)$ visitano punti diversi di A (agli istanti $0, 1, \ldots, m(\omega) - 1$), allora:

$$\zeta(s) = \prod_{\omega \in P} (1 - e^{-sm(\omega)})^{-1}$$

che spiega il nome per la analogia con la funzione zeta di Riemann, vista come funzione generatrice dei logaritmi dei numeri primi : $\zeta_{Riemann}(s) = \prod_p (1 - \exp -s \log p)^{-1}$. Di solito $\zeta(s)$ è definita per $\mathrm{Re}\, s$ grande, e dalle sue proprietà di analiticità si deducono informazioni sulla ditribuzione dei periodi delle orbite periodiche; proprio come dalle proprietà di analiticità si deducono informazioni sulla distribuzione dei numeri primi (v. Teorema dei Numeri Primi). Per i più semplici attrattori strani, v. Caos, la funzione zeta è olomorfa a destra di un certo $s_1 > 0$, tranne che per un polo semplice in $s_0 > s_1$ con residuo C: in tal caso segue la formula asintotica: $N_m \simeq mC\, e^{ms_0}$ (nel senso che il rapporto tende a 1 per $m \to \infty$ esponenzialmente); in particolare per m abbastanza grande tutti i valori di m sono periodi di qualche orbita periodica. La funzione zeta può essere definita anche per sistemi dinamici continui (A, S_t) come $\zeta(s) = \prod_{\omega \in P} (1 - e^{-sT(\omega)})^{-1}$ ove $T(\omega)$ è il periodo di ω e, nel caso dei più semplici sistemi caotici, ha notevoli proprietà di analiticità che permettono di trovare informazioni sulla distribuzione dei periodi delle orbite periodiche. Inoltre l' analogia

con la funzione zeta di Riemann è, in questi casi, assai più stretta perchè la ζ sopra definita verifica altre proprietà notevoli quali una relazione funzionale analoga a quella verificata dalla zeta di Riemann e, a volte, ha gli zeri posti su luoghi di punti semplici. Questo ha spinto alla ricerca, finora vana, di sistemi dinamici continui la cui funzione zeta sia la funzione zeta di Riemann.

Gas di sfere rigide, sistema di N particelle interagenti con una interazione di solo cuore duro (v.). Tale sistema è il più semplice modello di gas non perfetto e pare che già presenti fenomeni di transizione di fase (tipo liquido–gas), almeno se lo spazio ha tre dimensioni. Per quest' ultimo fenomeno si posseggono solo indicazioni desunte da esperimenti numerici.

Gas su reticolo: modello di gas in cui le particelle possono occupare solo posizioni che variano su un reticolo e, inoltre, ogni posizione può essere occupata al più da una sola particella. In questo modello le particelle si suppongono non mobili (e cioè dotate di energia cinetica nulla), ma distribuite con una distribuzione di probabilità microcanonica, canonica o gran canonica definite in termini della sola energia potenziale, che viene assegnata in ogni modello dando il potenziale di interazione (v. Interazione). I modelli di gas su reticolo, per quanto semplici, sono non banali e se la dimensione spaziale è > 1 possono presentare interessanti transizioni di fase tipo liquido gas o anche di altro tipo nei modelli più complessi. Il modello di Ising può essere interpretato come un modello di gas su reticolo se si interpreta che un dipolo $\sigma = +1$ significhi sito occupato e $\sigma = -1$ significhi sito vuoto. Così il modello di Ising (v.) ferromagnetico fornisce un interessante esempio di gas su reticolo con una transizione di fase liquido gas.

Gas perfetto, classico e quantico: sistema di particelle non mutuamente interagenti racchiuso in un contenitore a pareti elastiche perfette e in equilibrio termico con una riserva di calore (con la quale deve essere debolmente accoppiato. I gas perfetti classici hanno una equazione di stato verificante la legge dei gas perfetti; quelli quantici, invece, verificano la legge dei gas perfetti solo ad alta temperatura e bassa densità e altrimenti presentano fenomeni di degenerazione (v.).

Geometrie non euclidee: modelli di geometria piana, *e loro generalizzazioni*, che si ottengono sostituendo il piano euclideo E^2 con una superficie regolare bidimensionale Σ completa, immersa nello spazio tridimensionale E^3 o in E^d, $d \geq 3$ con metrica euclidea $\sqrt{x_1^2 + x_2^2 + \dots}$ oppure con metrica "pseudoeuclidea", $\sqrt{-x_1^2 + x_2^2 + x_3^2 + \dots}$ (in tal caso la superficie deve essere tale che il quadrato della lunghezza di ogni arco infinitesimo sia > 0), di particolare importanza nella teoria della relatività.
La nozione di punto diviene quella di punto di Σ. La nozione di retta diviene quella di *geodetica* (ossia di curva λ che minimizza la lunghezza dell'

arco di curva fra un suo qualunque punto $x \in \lambda$ ed un punto $y \in \lambda$ ad esso abbastanza vicino).
La nozione metrica di angolo in x fra due curve giacenti su Σ e passanti per x diviene quella di angolo fra i vettori tangenti alle due curve in x.
La condizione di completezza significa che ogni punto limite di punti della superficie è un punto *interno* alla superficie e implica anche l' *assenza di bordo*, cioè implica che dato un punto x ed un versore tangente $\underline{v}$ in x la (unica) geodetica che esce da x con direzione $\underline{v}$ ha lunghezza infinita nei due versi, ovvero è chiusa. Se la condizione di completezza non è valida, ma da ogni punto ed in ogni direzione esce una geodetica di lunghezza positiva in entrambe le direzioni, si dice che Σ è una porzione di superficie regolare.
Due porzioni di superficie regolare S, S' si dicono isomorfe, o "applicabili l' una sull' altra", se è possibile stabilire una corrispondenza biunivoca, bidifferenziabile ("diffeomorfismo globale") che conservi la lunghezza di archi di curva corrispondenti e gli angoli formati da coppie corrispondenti di archi uscenti da uno stesso punto.
La nozione di *movimento* di Σ diviene quella di isomorfismo fra Σ e se stessa. I movimenti formano un gruppo (che può essere banale e) che estende la nozione ordinaria di gruppo dei movimenti del piano (ossia gruppo delle rotazioni e traslazioni del piano).
Una regione semplicemente connessa delimitata da tre segmenti di geodetica si dice *triangolo* (o *triangolo geodetico*). Le seguenti tre proprietà caratterizzano il piano euclideo: 1) di essere una superficie globalmente diffeomorfa a E^2, 2) tale che, dati un punto A e due semirette da esso uscenti a, b ed un secondo punto A' e due semirette da esso uscenti a', b' formanti angoli in A e A' uguali, esiste sempre un movimento che trasporta A, a, b in A', a', b', e, 3) tale che la somma degli angoli interni di ogni triangolo (geodetico) Δ è π. Ovvero, 3'), tale che comunque siano dati una retta λ ed un punto x fuori di essa esiste una ed una sola retta parallela a λ e passante per x (due rette sono "parallele" se non hanno punti in comune).
In generale un triangolo geodetico Δ su una superficie Σ e con angoli al vertice α, β, γ è tale che:

$$\alpha + \beta + \gamma = \pi + \int_{\Delta} \rho(x) d\sigma(x)$$

ove $d\sigma$ è l' elemento d' area di Σ, in quanto superficie immersa in E^3 (o E^d): è il " teorema di Gauss" (che segue immediatamente dalla semplice osservazione che $\alpha + \beta + \gamma$ è una funzione additiva di area). : la quantità $\rho(x)$ è detta *curvatura gaussiana* in x ed ha la dimensione $[l^{-2}]$. Se Σ è immersa in E^3 ed è data da un' equazione parametrica: $x_3 = z(x_1, x_2)$ allora:

$$\rho(x) = \frac{z_{11} z_{22} - z_{12}^2}{(1 + z_1^2 + z_2^2)^2}, \qquad \text{se} \quad z_j \equiv \frac{\partial z}{\partial x_j},\ z_{ij} \equiv \frac{\partial^2 z}{\partial x_i \partial x_j}$$

e se Σ é immersa nello spazio pseudoeuclideo a tre dimensioni vale una formula analoga (ossia $\rho(x) = \frac{-z_{11}z_{22}+z_{12}^2}{(1-z_1^2-z_2^2)^2}$). Dunque se una superficie in E^3 ha curvatura > 0 allora deve delimitare una regione convessa, mentre se ha curvatura < 0 allora ogni suo punto è un punto di sella. Quindi in E^3 non possono esistere superfici limitate a curvatura negativa. Ma possono esistere immerse in spazi a dimensione maggiore. Assai più profonda è l' impossibilità di realizzare in E^d una superficie a curvatura costante negativa, globalmente diffeomorfa a E^2, *qualunque sia d* (teorema di Hilbert). In E^3 si possono però facilmente disegnare porzioni di superficie regolare a curvatura costante negativa. E superfici a curvatura costante negativa possono essere facilmente disegnate nello spazio pseudoeuclideo a tre dimensioni.

Un esempio di superficie a curvatura > 0 costante è la superficie della sfera di raggio R (che ha curvatura R^{-2}).

Un esempio, dovuto a Beltrami, di *porzione* di supeficie regolare a curvatura $-R^{-2} < 0$ costante è la superficie $\mathcal{T}$ di rivoluzione attorno all' asse z con curva meridiana data dalla *trattrice*: $z = R\log\frac{R+\sqrt{R^2-x^2}}{x} - \sqrt{R^2-x^2}$, $0 < x < R$ (la trattrice verifica l' equazione differenziale $\frac{dx}{dz} = -\frac{x}{\sqrt{R^2-x^2}}$, con $x = R$ per $z = 0$).

La geometria della superficie della sfera non verifica vari assiomi, della geometria euclidea (ad esempio non gli assiomi di incidenza né quelli di ordine: perchè le rette sono i cerchi massimi e quindi punti opposti sono congiunti da più rette, né sulle rette si può stabilire un ordinamento dei punti; verifica invece gli assiomi di congruenza, grazie al gruppo dei movimenti costituito dalle rotazioni, e quelli di continuità). Infine non verifica il postulato delle parallele.

Un esempio di superficie a curvatura negativa costante immersa nello spazio *pseudoeuclideo* a 3 dimensioni è la "pseudosfera": $x_1 = \sqrt{R^2 + x_2^2 + x_3^2}$ (che in termini di geometria ordinaria è un ramo di iperboloide non rigato) la cui curvatura è $-R^{-2}$ (come superficie dello spazio pseudoeuclideo). Sulla pseudosfera valgono tutti gli assiomi della geometria euclidea (inclusi quelli di congruenza grazie all' esistenza di un opportuno gruppo di movimenti a tre parametri) tranne quello delle parallele: è un "modello di geometria iperbolica". *Quindi* l' assioma delle parallele non è deducibile dagli altri. Il riconoscimento della possibilità che il postulato delle parallele non fosse deducibile dagli altri è dovuto a Gauss, Bolyai, Lobachevski e fu preceduto da un geniale quanto dimenticato libro di Saccheri, che però considerò questa conclusione "ripugnante" e ne dedusse, invece, la validità dell' assioma delle parallele. Gli mancò in realtà un modello concreto di geometria iperbolica che gli mostrasse la consistenza logica della geometria della quale aveva descritto le proprietà salienti: il modello fu realizzato molto dopo da Poincaré (la pseudosfera è un esempio ancora successivo).

La Fisica fornisce altri "modelli di geometria". Un esempio (equivalente al modello di Poincaré) assai importante è fornito dall' ottica geometrica. Dato un mezzo ottico Σ, consistente in una regione Σ del piano in cui l' indice di rifrazione nel punto x è $n(x)$ si può chiamare "punto" x un punto di Σ, "retta" la traiettoria di un raggio luminoso che passa per x, "angolo" fra due curve per x l' angolo fra le loro tangenti (nel piano) e "lunghezza" di un arco λ la sua lunghezza ottica: $\int_\lambda n\,ds$, ecc. Se tutti i raggi emergenti da un punto in una direzione arbitraria hanno lunghezza infinita si dirà che Σ è "senza bordo". La curvatura ora si definirà tramite la formula di Gauss. Se Σ è il semipiano superiore $y > 0$ e se $n(x,y)^2 = \frac{R^2}{y^2}$ allora la geometria di Σ generata dai raggi ottici è una geometria a curvatura costante negativa $-R^{-2}$ isomorfa alla geometria iperbolica sulla pseudosfera (e questo fu il modello di Poincarè).
La geometria appena definita verifica (quindi) tutti gli assiomi della geometria euclidea tranne quello delle parallele. Questa geometria ammette il gruppo dei movimenti definito dalle trasformazioni bilineari fratte del semipiano superiore (ossia posto $z = x + iy$ e $z' = x' + iy'$, la trasformazione g ha la forma $z' = \frac{az+c}{bz+d}$ con $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ matrice reale a determinante 1).
Altri importanti esempi di modelli di geometria sono generati dalla Meccanica Classica (via il collegamento con l' ottica geometrica fornito dal principio di Maupertuis). Ma gli esempi fra i piú notevoli per la Fisica sono forniti dalla teoria della Relativitá Generale di Einstein.
La unificazione dei metodi per definire modelli di geometrie sopra descritti è possibile e conduce alla nozione di varietà differenziabile con metrica riemaniana.
(*Bibliografia:* R. Bonola, *La geometria non euclidea*, Pavia, 1906; N. Efimov: *Higher geometry*, MIR, Mosca, 1980;

Ghiaccio, modello del ghiaccio di Pauling: si suppone che l' ossigeno si diponga su un cristallo perfetto (tipo wurtzite) in cui ogni atomo ha quattro primi vicini. Gli atomi di idrogeno si dipongono sui segmenti (legami) che connettono coppie di ossigeni primi vicini. Il modello suppone inoltre che su ogni legame si possa trovare un solo atomo di idrogeno e che questo possa occupare una sola fra due posizioni, l' una prossima ad un estremo e l' altra all' altro (*regola del ghiaccio*). In questo modello il valore dell' entropia residua del ghiaccio (v. Principio di Nernst) è S, se e^{SV} è il numero di modi di diporre gli atomi di idrogeno sui legami compatibilmente con la regola del ghiaccio. Il valore di S, nel limite $V \to \infty$ si calcola numericamente ed è interessante che nel modello analogo bidimensionale, in cui gli ossigeni occupano i siti di un reticolo quadrato il valore di S, si può calcolare esattamente (v. Modelli Risolubili).

Gibbs, distribuzione o stato di G.: in MSC è una distribuzione di probabilità sullo spazio delle fasi di un sistema, in equilibrio termodinamico, occupante l' intero spazio infinito. Uno stato di Gibbs è definito assegnando le distribuzioni di probabilità locali (v. funzioni di Correlazione). Cioè assegnando per ogni volume finito Λ le probabilità $f_\Lambda(x_1, \dots, x_n)\, dx_1 \dots dx_n\, /n!$ di trovare in Λ esattamente n particelle e di trovarle esattamente negli elementi di volume $dx_1, \dots, dx_n$ attorno ai punti $x_1, \dots, x_n$. Inoltre tali ditribuzioni locali devono verificare opportune relazioni che coinvolgono il potenziale interparticellare, la temperatura, etc; sono le equazioni DLR, (v. Limite termodinamico, Equazioni DLR). Alternativamente uno stato di Gibbs può anche essere definito dalla famiglia delle sue funzioni di correlazione, (v.), imponendo che esse verifichino le equazioni KS (v.). In MSQ uno stato di Gibbs è definito dalle matrici di densità ridotte (analoghe alle funzioni di correlazione classiche) imponendo che verifichino opportune equazioni (analoghe alle Equazioni di KS). Alternativamente si può definire uno stato di Gibbs imponendo la condizione KMS (v.). La teoria degli stati di Gibbs coincide con la teoria degli insiemi statistici e della loro equivalenza; e costituisce un quadro concettuale per la Meccanica Statistica analogo alla Meccanica Analitica per la Meccanica Classica, (v. Limite termodinamico).

Gibbs, potenziale di G.: è la funzione termodinamica $G = U + PV - TS$, (ove U= energia interna, P= pressione, V= volume, T= temperatura assoluta, S= entropia); il potenziale di Gibbs è legato in modo semplice al potenziale chimico μ: $\mu\rho V + (3/2)V \log(m\beta/2\pi) = G$, ove ρ è la densità e $\beta = 1/k_B T$ (talvolta si definisce il potenziale chimico via $\mu\rho V = G$, semplicemente); v. Termodinamica.

Goccia, modello a: modello per la teoria delle configurazioni microscopiche di una fase pura in una transizione liquido–gas, (v. Transizioni di fase, Contorni di Peierls).

Grad–Boltzmann, congettura, v. Boltzman–Grad.

Grammomolecola, quantità in grammi di una sostanza pura uguale al valore del peso molecolare delle molecole costituenti. Se la sostanza è un elemento monoatomico si dice anche grammoatomo, v. Avogadro.

Gran canonico, insieme statistico: v. Distribuzione gran canonica, IS.

Green–Kubo, formule: v. Coefficienti di Trasporto.

Gruppo di rinormalizzazione: metodo per la teoria delle fluttuazioni in sistemi anche assai diversi fra loro. Ha fornito la prima teoria generale dei fenomeni critici, per sistemi governati da forze a corta portata, alternativa a quella di campo medio (v., che li descrive in modo insoddisfacente). Nella Meccanica Statistica è, sotto certi aspetti, un raffinamento della teoria di campo medio cui essenzialmente si riduce nel caso in cui si considerino sistemi in uno spazio a dimensione d abbastanza alta (il valore esatto dipende

dal sistema in esame, ad esempio per i punti critici di transizioni liquido gas ovvero di sistemi con transizioni ferromagnetiche d deve essere ≥ 4). In altra forma è applicabile anche allo studio della Teoria dei campi quantizzati, perchè in essa l' oggetto di maggiore interesse è proprio la teoria delle fluttuazioni del vuoto. Ha condotto ad una chiarificazione della teoria della rinormalizzazione (v.) e da questo ha ricevuto il nome, (v. Teoria Quantistica dei Campi, Fenomeni Critici). Come la teoria di campo medio il metodo del gruppo di rinormalizzazione, dovuto nella sua forma finale a K. Wilson, è nei casi più interessanti una teoria approssimata; ma ha il vantaggio di essere concettualmente semplice e di fornire risultati diversi dalla teoria di campo medio e più conformi alle osservazioni sperimentali (nei casi in cui differisce da essa).

H, teorema: teorema di irreversibilità per le soluzioni dell' equazione di Boltzmann; H è eta maiuscola ed è il simbolo usato da Boltzmann per denotare l' entropia; v. Teorie Cinetiche, MSC.

Heisenberg, modello di H.: modello per il ferromagnetismo. Il sistema è pensato come un cristallo ideale formante un reticolo immobile classico e in ogni vertice del reticolo è posto un dipolo magnetico quantico i cui stati sono descritti dai vettori di uno spazio bidimensionale. La hamiltoniana del sistema è espressa in termini delle osservabili $\underline{S}^i = (\sigma_x^i, \sigma_y^i, \sigma_z^i)$, ove $\sigma_x, \sigma_y, \sigma_z$ denotano le matrici di Pauli, degli spin associati ai vari siti i del cristallo. Si suppone:

$$H = -\sum_{\langle i,j \rangle} J\underline{S}^i \cdot \underline{S}^j - \sum_i h\sigma_z^i$$

ove $\langle i, j \rangle$ significa che la somma è estesa solo alle coppie i, j di siti reticolari che sono primi vicini nel cristallo; $J > 0$, h è il campo magnetico supposto parallelo all' asse z. Il modello può essere esteso a sistemi con dipoli a spin più elevato di $1/2$ e a interazioni coinvolgenti altre coppie di siti reticolari oltre i primi vicini. Se $J < 0$ si ha il M.H. antiferromagnetico. Si hanno indicazioni numeriche e congetture teoriche che il M.H. presenta una transizione di fase ferromagnetica a temperatura abbastanza bassa e campo magnetico nullo (se la dimensione dello spazio è $d \geq 3$). Anche il modello antiferromagnetico ha transizioni di fase (antiferromagnetiche, v. accoppiamento antiferromagnetico) a bassa temperatura e questo risultato è stato mostrato in tutto rigore matematico (a differenza del caso ferromagnetico) da F. Dyson, E. Lieb, B. Simon. Il M.H. ha un analogo classico: il modello di Ising–Lenz, (v.).

Hilbert, metodo di H. per l' equazione di Boltzmann: si suppone che la funzione di distribuzione $f(\underline{x}, \underline{v})$ incognita dell' equazione di Boltzmann ammetta uno sviluppo in serie in un parametro ε, che viene posto a coefficiente del termine di collisione nella forma $1/\varepsilon$ e, alla fine dei calcoli, vien posto

uguale ad 1. Si impone che l' equazione sia soddisfatta ordine per ordine. Si ottengono infinite equazioni, la prima delle quali esprime che la f all' ordine 0 deve essere una maxwelliana $n_0(\underline{r})e^{-\beta_0(\underline{r})m(\underline{v}-\underline{u}_0(\underline{r})^2/2}(m\beta_0(\underline{r})/2\pi)^{3/2}$ ove $n_0(\underline{r}), \underline{u}_0(\underline{r}), \beta_0(\underline{r})$ sono tre funzioni che si interpretano come la densità in $\underline{r}$, la velocità media in $\underline{r}$ e la temperatura assoluta $T_0(\underline{r}) = \beta_0(\underline{r})^{-1}/k_B$ in $\underline{r}$ (se k_B è la costante di Boltzmann); m denota la massa delle molecole del gas. L' equazione del primo ordine impone (per poter essere risolubile) che le grandezze *idrodinamiche* $n_0(\underline{r}), \underline{u}_0(\underline{r}), \beta_0(\underline{r})$ verifichino le equazioni di Eulero comprimibili linearizzate (v.) della gasdinamica. La sua soluzione può essere espressa in termini di cinque funzioni arbitrarie $n_1(\underline{r}), \underline{u}_1(\underline{r}), \beta_1(\underline{r})$, che hanno il significato fisico di termini del primo ordine in uno sviluppo in serie di ε delle grandezze dell' idrodinamica $n(\underline{r}) = n_0(\underline{r}) + \varepsilon n_1(\underline{r}) + \ldots$, $\underline{u}(\underline{r}) = \underline{u}_0(\underline{r}) + \varepsilon \underline{u}_1(\underline{r}) + \ldots$, $\beta(\underline{r}) = \beta_0(\underline{r}) + \varepsilon\beta_1(\underline{r}) + \ldots$; a loro volta queste funzioni devono verificare un' equazione lineare affinchè la equazione del secondo ordine sia risolubile, etc.. Il metodo di Hilbert fa dunque vedere in quale modo si possa tentare una connessione fra le equazioni macroscopiche della Meccanica dei Fluidi (v.) e l' equazione di Boltzmann, tuttavia è poco chiaro in che senso lo sviluppo di Hilbert approssimi le soluzioni dell' equazione di Boltzmann, e conseguentemente è poco chiara la connessione fra l' equazione di Boltzmann e le equazioni della fluidodinamica macroscopica. Si ritiene che il metodo sia applicabile quando si considerano particolari condizioni iniziali e in situazioni in cui il rapporto fra il tempo t_{lcm}, di percorrenza di una distanza pari al libero cammino medio da una molecola con velocità media v è molto piccolo rispetto al tempo necessario all' evoluzione idrodinamica τ, (τ ha lo stesso ordine di grandezza di L/v) e il rapporto $(t_{lcm}/\tau)^k$ viene identificato con l' ordine di grandezza del termine f_k dello sviluppo di H. (v. Teorie Cinetiche, Coefficienti di Trasporto, metodo di Chapman Enskog, Gerarchia di BBGKY, Equazioni di Eulero, Equazioni di Navier Stokes).

Inerzia, assi e momenti principali di: dato un sistema di n punti materiali soggetto al vincolo di rigidità, v. *sistemi rigidi*, si dice che $\vec{\imath}_1, \vec{\imath}_2, \vec{\imath}_3$ sono assi principali d' inerzia in un punto O del sistema o solidale con esso se, dette m_1 la massa dell' i-mo punto e x_i, y_i, z_i le coordinate cartesiane dello stesso punto nel riferimento *solidale* $(O; \vec{\imath}_1, \vec{\imath}_2, \vec{\imath}_3)$ si ha:

$$\sum_i m_i x_i y_i = \sum_i m_i y_i z_i = \sum_i m_i z_i x_i = 0$$

e le quantità $I_1 = \sum_i m_i(y_i^2 + z_i^2)$, $I_2 = \sum_i m_i(z_i^2 + x_i^2)$, $I_3 = \sum_i m_i(x_i^2 + y_i^2)$, si dicono *momenti principali d' inerzia*.

Se $I_i \neq I_j$, per $i \neq j$, allora gli assi principali sono univocamente determinati a meno di permutazioni dei nomi e cambiamenti di verso; inoltre non dipendono dal punto O. Se due momenti sono uguali e il terzo diverso

da essi allora solo l' asse corrispondente a quest' ultimo è determinato univocamente: gli altri due possono essere scelti arbitrariamente nel piano ortogonale e, in questo caso, il solido si dice di *tipo cilindrico* attorno a O. Se infine i tre momenti sono uguali il solido si dice di *tipo sferico* attorno a O e gli assi principali possono essere scelti arbitariamente.

Particolare importanza hanno i momenti d' inerzia *baricentrici*, ossia relativi al baricentro del sistema: se li denotiamo I_1^0, I_2^0, I_3^0 e se ξ, η, ζ sono le coordinate di un altro punto O in $(G; \vec{\imath}_1, \vec{\imath}_2, \vec{\imath}_3)$ allora i momenti d' inerzia relativi a O sono semplicemente $I_1 = I_1^0 + M(\eta^2 + \zeta^2)$, $I_2 = I_2^0 + M(\zeta^2 + \xi^2)$, $I_3 = I_3^0 + M(\xi^2 + \eta^2)$.

La posizione, in un riferimento solidale qualsiasi $(O; \vec{\imath}, \vec{\jmath}, \vec{k})$, degli assi principali si calcola a partire dalla *matrice d' inerzia*: se x_1^i, x_2^i, x_3^i sono le coordinate di un generico punto $\underline{x}^i$ in $(O; \vec{\imath}, \vec{\jmath}, \vec{k})$ e $\delta_{\alpha\beta}$ è la matrice identità 3×3, la matrice d' inerzia (3×3), in questo riferimento, è definita da:

$$I_{\alpha\beta} = \sum_i m_i (x_\alpha^i x_\beta^i - \delta_{\alpha\beta} (\underline{x}^i)^2)$$

con $\alpha, \beta = 1, 2, 3$. Gli assi principali sono allora tre autovettori ortonornali di questa matrice e i rispettivi momenti d' inerzia principali sono i corrispondenti autovalori.

Interazione a due corpi: modello di forza molecolare in cui la energia potenziale di interazione si può esprimere come somma delle energie potenziali di interazione fra tutte le coppie di molecole. Ad esempio si consideri un sistema di molecole puntiformi identiche e si supponga che l' energia potenziale totale nella configurazione in cui le molecole occupano i siti $x_1, ..., x_N$ sia, per una scelta opportuna di φ:

$$V(x_1, \ldots, x_N) = \sum_{i<j} \varphi(x_i - x_j)$$

Talvolta si dice che un potenziale di questa forma verifica la proprietà di *additività*. In MS si considerano a volte modelli di interazione in cui la energia potenziale contiene oltre a un contributo della forma precedente anche termini *a tre corpi*, come:

$$V'(x_1, ..., x_N) = \sum_{i<j<k} \varphi'(x_i, x_j, x_k)$$

o a *quattro o più* corpi. Le funzioni $\varphi, \varphi' \ldots$ si dicono rispettivamente potenziali a due, tre... corpi. Le leggi di interazione a due corpi hanno la proprietà che la forza che si esercita su una delle molecole si può interpretare come risultante delle forze dovute a ciascuna delle altre molecole. Se l' interazione contiene termini a tre (o più) corpi questa interpretazione non

è più possibile. Se i potenziali $\varphi, \varphi', ...$ si annullano quando la distanza massima r fra i loro argomenti supera un valore r_0 si dice che l' interazione ha *portata finita* r_0; questa locuzione si continua ad usare se i potenziali anzichè annullarsi per $r > r_0$ si annullano esponenzialmente per $r \to \infty$ su scala r_0, cioè $\propto e^{-r/r_0}$.

Ipotesi ergodica, di Boltzmann, (v. Boltzmann).

Ising, modello di E. Ising–W. Lenz: modello per il ferromagnetismo. Il sistema è pensato come un cristallo ideale formante un reticolo immobile classico e in ogni vertice i del reticolo è posto un dipolo magnetico i cui stati sono descritti dal valore $\sigma_i = \pm 1$, dell' orientazione del dipolo (che è suscettibile di assumere solo la configurazione in cui è parallelo o antiparallelo ad una data direzione prefissata). La energia delle configurazioni del sistema è espressa da:

$$H = -\sum_{\langle i,j \rangle} J\sigma_i\sigma_j - \sum_i h\sigma_i$$

ove $\langle i, j \rangle$ significa che la somma è estesa solo alle coppie i, j di siti reticolari che sono primi vicini nel cristallo; $J > 0$, h è il campo magnetico. Il modello può essere esteso a sistemi con dipoli a spin più elevato di $1/2$ e a interazioni coinvolgenti altre coppie di siti reticolari oltre i primi vicini. Può inoltre essere reinterpretato come un modello per la MS di un gas (detto gas su reticolo, v.). Se $J < 0$ si ha il M.I. antiferromagnetico. Il M.I. può essere esattamente risolto se la dimensione d dello spazio è $d = 1$ o, se si suppone anche $h = 0$, $d = 2$, e il reticolo è quadrato (o di alcune altre forme semplici). La soluzione nel caso $d = 1$ è dovuta a Ising, mentre nel caso $d = 2$ è stata scoperta da L. Onsager nel 1942-48 (v. Modelli Risolubili). Nel caso $d = 1$ il modello, come tutti i modelli unidimensionali con interazioni a corta portata, non presenta transizioni di fase, (v. Landau, argomento unidimensionale), e quindi il modello è, in un certo senso, poco interessante. Diversa è l' importanza del caso $d = 2$: in questo caso già Peierls e van der Waerden avevano dimostrato l' esistenza di una transizione di fase se $d \geq 2$ e la simmetria di dualità (v.) permetteva di conoscere il valore esatto della temperatura critica (nella ipotesi che fosse unica);e l' interesse di questo risultato fu di chiarire una volta per tutte che la MSC poteva prevedere il fenomeno delle transizioni di fase. La soluzione esatta di Onsager, al di là del suo interesse matematico, mostrò inoltre che i modelli della Meccanica Statistica con interazioni a corta portata potevano dare luogo a transizioni di fase con esponenti critici molto diversi da quelli previsti dalla teoria di campo medio. Alcuni esponenti critici del modello di Ising sono noti a 2 dimensioni, taluni rigorosamente, altri sulla base di argomenti euristici abbastanza solidi; nessuno è noto rigorosamente a 3 dimensioni (sebbene si creda di conoscere gli esponenti critici con grande

approssimazione grazie a esperimenti numerici, v.); se la dimensione è ≥ 4 molti esponenti critici sono noti (Wilson, Aizenman) e coincidono con quelli previsti dalla teoria di campo medio (v. Teoria di Campo Medio, Gruppo di Rinormalizzazione). Quest' ultima proprietà di banalità del M.I., al punto critico, ha indotto a congetturare l' impossibilità delle teorie di campo puramente scalari nella teoria relativistica dei campi (se la dimensione dello spazio tempo è 4): la base di questa congettura è la connessione fra la teoria dei campi e la meccanica statistica classica messa in luce da K. Wilson con il metodo del gruppo di rinormalizzazione (v.).

Invarianza di scala, proprietà di un sistema in equilibrio termodinamico le cui funzioni di correlazione sono asintoticamente, a grande distanza, funzioni omogenee. Questa proprietà è di solito caratteristica di un sistema in equilibrio al punto critico. Ad esempio in una transizione liquido gas le funzioni di correlazione spaziale a 2 punti ($\rho(\underline{r}_1, \underline{r}_2)$, v. Correlazione), dovrebbero avere, al punto critico, grado di omogeneità $\omega_2 = d - 2 + \eta$, e questa proprietà definisce l' indice critico η.

Irreversibilità: impossibilità che un sistema a molte particelle percorra a ritroso la sua evoluzione. Fallace è dire che basta invertire ad un dato istante tutte le velocità delle particelle: è evidentemente impossibile eseguire questa operazione in un tempo che resti finito all' aumentare della precisione. D' altra parte il carattere instabile dei moti hamiltoniani (v. Caos) richiede che la precisione con cui si devono fissare i dati iniziali (cioè posizioni uguali e velocità opposte a valori assegnati) per osservare il moto a ritroso svolgersi per un tempo prefissato, deve essere una precisione che cresce esponenzialmente con il numero di particelle. Fallace sarebbe anche dire, come alcuni critici di Boltzmann, che se si attendesse abbastanza si osserverebbe il sistema ripercorrere spontaneamente a ritroso la sua evoluzione: in base all' ipotesi ergodica Boltzmann stimò tale tempo e trovò che il suo ordine di grandezza doveva essere di gran lunga più grande dell' età dell' universo già per sistemi con qualche decina di particelle (v. ipotesi ergodica di Boltzmann, Demonietto di Maxwell, Ricorrenza).

Kac, equazione di stato con potenziale di K.: si considera un modello di gas generato da un potenziale di interazione fra particelle avente la forma: $\varphi(r) = \gamma^d V(\gamma r) + V_a(r)$ ove γ è un parametro, V è una funzione decrescente rapidamente all' ∞ e V_a è un potenziale a cuore duro di raggio $a > 0$, (v.). Se $P = F_\gamma(\rho, T)$ è l' equazione di stato per questo gas l' equazione di stato con potenziale di Kac è per definizione $P = F(\rho, T)$ con $F = \lim_{\gamma \to 0} F_\gamma$. Se V è attrattivo e se $d = 1$ si dimostra (M. Kac, G. Uhlenbeck, P. Hemmer, 1963) che tale equazione di stato coincide con la equazione di van der Waals corredata della regola di Maxwell (v.); se $d > 1$ si dimostra (J. Lebovitz, O. Penrose, 1966) che posto $A = 2^{-1} \int V(r) d^d \underline{r}$ e se $P = F_a(\rho, T)$ è l' equazione di stato per l' interazione di solo cuore duro allora l' equazione

di stato di un modello con potenziale di Kac è $P+A\rho^2 = F_a(\rho, T)$ corredata della regola di Maxwell. Dunque la teoria dell' equazione di stato con potenziale di Kac fornisce una analisi rigorosa delle condizioni di validità dell' equazione di van der Waals, (v. Transizioni di Fase, Serie e Teorema del Viriale).

KMS , condizione: in un sistema quantico in equilibrio termodinamico si considerano due qualsiasi grandezze osservabili descritte da due operatori A e B. Se $A(t)$ è la osservabile in cui evolve A nel tempo t per l' evoluzione generata dalla hamiltoniana H del sistema, si considera la funzione $F(t; A, B) = \langle A(t)B\rangle$, valore medio del prodotto $A(t)B$ nello stato di equilibrio a temperatura T (cioè $F(t) = (\mathrm{Tr}\, e^{-\beta H}\, A(t)B)/(\mathrm{Tr}\, e^{-\beta H})$) con $\beta = 1/k_B T$. La condizione KMS allora è $F(t; A, B) = F(t + i\beta; B, A)$ (se $\hbar = 1$) ed è una immediata conseguenza della ciclicità della traccia. Questa condizione non è banale, perchè per essere veramente scritta richiede la conoscenza di molte proprietà dell' operatore H, energia del sistema. Può essere usata per generare le equazioni cui devono soddisfare le funzioni di correlazione del sistema in equilibrio; è stata anche usata per mostrare importanti disuguaglianze. Che hanno portato, ad esempio, ad escludere la possibilità di transizioni di fase in sistemi bidimensionali con forze a corta portata e con hamiltoniana invariante rispetto ad un gruppo di simmetria continuo (teorema di Mermin e Wagner). Tali sistemi non possono presentare transizioni di fase dovute a rottura spontanea di simmetria (v.). Così la simmetria di traslazione non può essere spontaneamente rotta in un sistema bidimensionale con forze a corta portata, escludendo quindi l' esistenza di stati cristallini come stati di equilibrio termodinamico di questi sistemi; cosi' anche il modello di Heisenberg (v.) a 2 dimensioni non può presentatre stati con magnetizzazione spontanea. KMS sta per Kubo, Martin, Schwinger.

Landau: argomento unidimensionale di L.: argomento generale che mostra l' impossibilità di transizioni di fase a temperatura $T > 0$ in sistemi unidimensionali con interazioni a corta portata. Lo stesso tipo di analisi in sistemi a dimensione $d \geq 2$ porta, invece, alla dimostrazione della possibilità di transizioni di fase, almeno in sistemi semplici e in cui la transizione si manifesta come una roottura spontanea di simmetria (v. Transizioni di f Fase, Argomento di Peierls). Una analisi matematica rigorosa in vari sistemi unidimensionali fu svolta (1949) da L. van Hove.

Lanford, teorema di O.L.: si considera un sistema di sfere rigide con densità n e raggio a. Si suppone che nello stato iniziale i centri delle sfere siano distribuiti a caso con funzioni di correlazione che fattorizzano e che sono vicine ad una distribuzione maxwelliana per quel che riguarda la distribuzione degli impulsi. Si fa evolvere il sistema per un tempo t secondo l' evoluzione hamiltoniana (supponendo gli urti perfettamente elastici).

Fissato $t > 0$ si considera il limite della funzione di correlazione a una particella $\rho_{n,a}(\underline{p}, \underline{r}; t)$ per $n \to \infty$, $a \to 0$ in modo che $na^2 = cost$: si dimostra che il limite $\lim n^{-1}\rho_{n,a}(\underline{p}, \underline{r}; t) = f(\underline{p}, \underline{r}; t)$ esiste e risolve l' equazione di Boltmann; inoltre le funzioni di correlazione a k punti, divise per n^k sono semplicemente uguali al prodotto delle corrispondenti funzioni f a un punto (propagazione del caos). Questo teorema mostra rigorosamente la compatibilità fra un' evoluzione microscopica hamiltoniana e reversibile con una dinamica macroscopica irreversibile. La limitazione più grossa è che la sua validità è stata dimostrata solo per t assai piccolo (circa il 20% del tempo medio fra due urti). Il risultato fu congetturato in forma precisa da H. Grad nel 1958 e va sotto il nome di congettura di Grad–Boltzmann, (v. Teorie Cinetiche, MSC).

Langevin, equazione di L.: v. Moto Browniano.

Lee–Yang, teorema di L.Y.: stabilisce che la funzione di partizione del modello di Ising attrattivo (con potenziale anche non di primo vicino) pensata come funzione della variabile $z = e^{\beta h}$ non può annullarsi che per valori complessi e di modulo 1 di z. Da questo teorema L.Y. dedussero che il modello di Ising poteva avere più fasi in equilibrio solo a campo magnetico nullo e che il fenomeno matematico che era all' origine della transizione doveva essere che, per un valore critico di β e per tutti i valori più grandi, l' insieme degli zeri della funzione di partizione si avvicinava sempre più al punto $z = 1$ al crescere del volume del sistema; mentre per valori piccoli di β (ossia a temperatura alta) gli zeri della funzione di partizione si mantenevano ad una distanza positiva dall' asse reale. Gli zeri della funzione di partizione sono connessi alle singolarità delle funzioni termodinamiche semplicemente perchè queste sono espresse in termini del logaritmo della funzione di partizione. A tutt' oggi però non si è riusciti a determinare la densità limite degli zeri della funzione di partizione sul cerchio unitario della variabile complessa z (che equivarrebbe, come mostrarono L.Y., alla soluzione esatta del modello di Ising in campo magnetico non nullo), (v. Transizioni di fase).

Legge di Planck: in un corpo nero (v.) in equilibrio termico alla temperatura T la densità di energia u_ν della radiazione con frequenza fra ν e $\nu + d\nu$ è data da:

$$u_\nu = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}$$

ove h, c, k_B sono rispettivamente la costante di Planck, la velocità della luce e la costante di Boltzmann, v. Statistiche Quantiche, MSC.

Leggi di scala, per gli esponenti critici: sono relazioni che legano gli esponenti critici $\alpha, \beta, \gamma, \delta, \nu, \mu$ fra di loro (si veda la voce Esponenti Critici per la definizione di queste quantità). Si dividono in relazioni che non coinvolgono la dimensione d dello spazio e nelle altre, che sono chiamate leggi

di iperscala. Esempi di leggi del primo tipo, in riferimento alla transizione liquido gas, sono:

$$\begin{aligned} \alpha =& 2 - \beta(\delta + 1) \\ \gamma =& \beta(\delta - 1) \\ \alpha + 2\beta + \gamma =& 2 \\ \gamma =& (2 - \eta)\nu \end{aligned}$$

Esempi di leggi di iperscala, sempre in riferimento alla stessa transizione e nell' ipotesi di forze a corta portata e $d \leq 4$ sono:

$$\begin{aligned} d\nu =& 2 - \alpha \\ 2 - \eta =& d(\delta - 1)/(\delta + 1) \\ \mu =& (d - 1)\nu \end{aligned}$$

Le leggi di iperscala non sono valide se $d \geq 4$. (v. Punto Critico, Esponenti Critici, Fenomeni Critici, Transizioni di Fase).

Legge di Stefan-Boltzmann: v. Costante di S.B..

Leggi di iperscala: v. Leggi di scala.

Lennard Jones , potenziale di Lennard Jones J.: modello empirico di potenziale di interazione (v.) fra molecole. E' un potenziale a due corpi additivo caratterizzato da due parametri, la portata σ e la intensità ε:

$$\varphi(r) = 4\varepsilon((\frac{\sigma}{r})^n - (\frac{\sigma}{r})^m)$$

e m, n sono due ulteriori parametri che determinano la repulsione all' origine e la coda all' infinito. Il modello con $m = 6, n = 12$ è considerato un buon modello per l' interazione fra i gas rarefatti. Lennard Jones (1924) calcolò il secondo coefficiente del viriale (per $m = 4$ e $n = 8, 10, 40/3$). Il modello ha ricevuto un fondamento teorico dal lavoro di Wang (1927): in questa, che fu una delle prime applicazioni della Meccanica Quantica, si mostrò che il potenziale fra due atomi di idrogeno a grande distanza doveva decadere come r^{-6}; il valore $n = 12$ invece ha carattere empirico (v. Serie del Viriale e per la bibliografia, v. Esponenti Critici).

Limite termodinamico: studio delle proprietà statistiche di un sistema di particelle nel limite in cui gli si lascia occupare tutto lo spazio mantenendo fisse due variabili (o più nei sistemi non monomolecolari o con gradi di libertà interni) atte a determinare lo stato di equilibrio termodinamico del sistema (ad esempio, la densità numerica e la temperatura, o la densità numerica e la densità di energia).

Liouville, teorema di L.: se si fanno evolvere tutti i punti di una regione $\underline{\omega}$ dello spazio delle fasi di un sistema meccanico hamiltoniano, per un tempo t prefissato, il volume di $\underline{\omega}$ resta inalterato (anche se $\underline{\omega}$ si deforma grandemente), v. Meccanica Classica, MSC.

Localizzazione di elettroni: gli elettroni di conduzione in un metallo perfetto sono, in prima approssimazione, pensabili come elettroni non interagenti mutuamente e posti in un campo esterno periodico. In questo caso lo spettro dell' operatore hamiltoniano è puramente continuo e non possiede autostati normalizzabili. Se il cristallo non è perfetto è possibile che l' operatore hamiltoniano possieda stati normalizzabili detti stati localizzati. Questo può avvenire se il cristallo è impuro con impurità distribuite a caso. Si consideri, ad esempio, il caso in cui le impurità sono schematizzabili pensando che il potenziale periodico generato dal reticolo in ogni cella è sostituito da un potenziale che assume valori casuali, e indipendenti in ogni cella. Allora si può mostrare che esistono stati localizzati almeno se la dispersione del potenziale stocastico in ogni sito è abbastanza grande; esiste però, se la dimensione dello spazio è maggiore di 1, un valore dell' energia (soglia di mobilità) tale che non possono esistere stati localizzati con energia superiore ad esso (se la dimensione è 1 tutti gli stati elettronici sono, invece, localizzati).

Lorentz, modello di L.: è un modello in cui si considera una particella puntiforme che si muove liberamente in uno spazio sparso di ostacoli sferici di raggio a distribuiti a caso: quando la particela urta un ostacolo viene riflessa elasticamente. Se i centri degli ostacoli sono ditribuiti a caso con distribuzione di Poisson (v.) con densità n si può dimostrare che il moto della particella è descritto, nel limite $n \to \infty, a \to 0$ con $na^2 = cost$, dall' equazione di Boltzmann lineare che si può dedurre sulla base di argomenti euristici. Fu questo il primo caso in cui fu mostrata la validità della congettura di Grad–Boltzmann (v.) (anche se per un modello con equazione di Boltzmann lineare). In questo caso la limitazione fondamentale, ancora non superata nel caso del gas di sfere rigide, che restringe la dimostrazione della congettura di Grad–Boltzmann a tempi piccoli rispetto al tempo medio fra due collisioni *non* è necessaria e la congettura è mostrata valida per tutti i tempi, (v. teorema di Lanford). Una delle questioni più interessanti sul modello di L. è se il moto della particella è un moto di diffusione (v.). Si crede che ciò sia vero se la dimensione dello spazio è 3 ma non se è 2. E' noto (Cohen, 1961) che se si tenta di esprimere il coefficiente di diffusione D in una serie di potenze nella densità degli ostacoli si trova che nD ha coefficiente di ordine 1 divergente se $d = 2$ mentre se $d = 3$ il coefficiente di ordine 1 è finito ma quello di ordine 2 diverge (si veda v. Coefficienti di Trasporto per una discussione dell' interpretazione in termini della diffusione di questo notevole fatto), v. BBGKY.

Lorenz, modello di E. Lorenz: modello originato dalla teoria delle previsioni metereologiche che illustra la possibilità di moti caotici, v. Caos.

Lunghezza di Correlazione: in un sistema omogeneo descrivente un liquido o un gas con densità ρ, in equilibrio termodinamico, si definisce la

funzione di correlazione spaziale troncata a due punti (v. Correlazione) come $h(r) = \rho(\underline{r}_1, \underline{r}_2) - \rho^2$, con $r = |\underline{r}_1 - \underline{r}_2|$. Nelle fasi pure questa funzione tende a 0 per $r \to \infty$. Se tende a zero esponenzialmente su una scala di lunghezza κ^{-1} allora κ^{-1} definisce la L.C.. Nei sistemi con interazione a corta portata κ^{-1} è $< +\infty$ tranne che al punto critico, che può essere pertanto definito come il punto dello spazio dei parametri termodinamici in cui la L.C. diverge. La nozione di lunghezza di correlazione si estende anche a transizioni di fase diverse da quella liquido gas, v. Esponenti Critici, Fenomeni Critici, Punto Critico.

Magnetizzazione spontanea, in un sistema di particelle dotate di dipolo magnetico e con interazione invariante per rotazione simultanea di tutti i dipoli può accadere che esistano stati di equilibrio termodinamico con magnetizzazione media non nulla, in assenza di campo magnetico esterno: si dice allora che lo stato di equilibrio in questione presenta magnetizzazione spontanea. Il fenomeno corrisponde ad una delle più semplici transizioni di fase (v. Transizioni di Fase, Rottura Spontanea di Simmetria, Ferromagnetismo, Modello di Ising, Modello di Heisenberg).

Matrici di densità , ridotte: è la nozione che in meccanica statistica quantistica sostituisce le funzioni di correlazione (v.). In un sistema di N particelle identiche chiuse in un volume V e descritte da un hamiltoniana H le MDR a n punti sono funzioni di $2n$ argomenti: e si esprimono facilmente in termini degli operatori di creazione e annichilazione (v.) $a_z^{\pm}$ come:

$$\rho(\underline{x}_1, \ldots, \underline{x}_n; \underline{y}_1, \ldots, \underline{y}_n) = \frac{\mathrm{Tr} e^{-\beta H} a^+_{x_1} \ldots a^+_{x_n} a^-_{y_1} \ldots a^-_{y_n}}{\mathrm{Tr} e^{-\beta H}}$$

Queste funzioni verificano equazioni analoghe alle equazioni KS (v.) della MSC, (J. Ginibre (1965)). Anche le funzioni di correlazione locali della MSC hanno un analogo quantico.

Maxwell, costruzione o regola di: è una regola che si può seguire quando si ha a disposizione una equazione di stato approssimata per migliorarne la precisione. Se l' equazione di stato prevede che ad una certa temperatura T la isoterma abbia una regione in cui la comprimibilità è negativa allora la regola di Maxwell richiede che si tracci, sul grafico p, ρ^{-1} della isoterma in questione una linea orizzontale con estremi sulla isoterma stessa e ad una altezza tale che l' area, con segno, della curva chiusa formata dal segmento orizzontale e dalla parte di isoterma che si appoggia ai suoi estremi sia nulla. Si sostituisce poi l' isoterma con la curva in cui il segmento orizzontale rimpiazza la parte dell' isoterma che si appoggia su di essa. Si ottengono cosi' nuove curve (monotone crescenti, e quindi mai a derivata negativa) che sono chiamate le isoterme della equazione di stato corretta dalla *costruzione di Maxwell*. In certe situazioni limite si può mostrare che questa costruzione conduce alla equazione di stato esatta (v. Potenziale di Kac).

Maxwell , demonietto di M.: mitico essere in grado di eseguire, su un sistema, misure di precisione arbitraria in tempi arbitrariamente piccoli e producendo trascurabili perturbazioni allo stato microscopico del sistema. La sua più celebre impresa è di avere impedito il passaggio di molecole da una metà di un contenitore all‘ altra, pur permettendo il passaggio nell‘ altro senso, attraverso un pertugio di dimensioni atomiche praticato in una parete divisoria, realizzando cosi' la inversione di una trasformazione irreversibile. Dedicò una simile impresa alla simultanea inversione delle velocità delle particelle di un gas, lasciando inalterate le loro posizioni. Gli strani fenomeni che il demonietto è cosi' in grado di produrre contraddicono il secondo principio della Termodinamica e, quindi, non sono osservabili: nel secondo caso, ad esempio, perchè, come dimostrato da Boltzmann, il tempo necessario perchè si verifichino spontaneamente, e quindi un essere umano possa osservarli senza l‘ aiuto del demonietto, sono di gran lunga superiori all‘ età dell‘ universo, (v. Boltzmann, ipotesi ergodica, Ricorrenza).

Maxwell–Boltzmann , distribuzione di M.B.: v. Distribuzione delle velocità.

Mermin–Wagner,teorema di M.W.: v. KMS

Mescolamento, proprietà di un sistema dinamico metrico $(\underline{\omega}, S, \mu)$ (v.); se per ogni coppia f, g di funzioni (μ–misurabili) su $\underline{\omega}$ si ha

$$\lim_{n\to\infty} \int f(S^n x) g(x) \mu(dx) = (\int f(x)\mu(dx))(\int g(x)\mu(dx))$$

allora il sistema si dice mescolante. La proprietà di mescolamento è più forte della proprietà di ergodicità (v. Sistemi Dinamici).

Microcanonico, insieme statistico (v. Distribuzione microcanonica, IS).

Modello a Goccia, v. Goccia

Modello Risolubile, in MSC: è un modello di un sistema con N particelle descritto da una interazione talmente spaciale (non necessariamente semplice) da permettere il calcolo esatto, senza approssimazioni, di *qualche* grandezza termodinamica associata a *qualche* stato di equilibrio termodinamico, (v. Modelli Risolubili in MS).

Monodo: termine usato da Boltzmann per denotare quello che oggi si chiama un insieme statistico (v.). L‘ abbreviazione (di Boltzmann stesso) di ergomonodo (M. caratterizzato dal valore dell‘ energia) in ergodo (v.), per descrivere l‘ insieme microcanonico, ha dato origine all‘ aggettivo ergodico. E la teoria ergodica ha questo nome perchè natag per giustificare l‘ uso degli ergodi per la descrizione degli stati di equilibrio termodinamico. Etimologicamente M. deriva da μόνος(unico) e εἶδος (aspetto), forse connesso ai monodi platonici o/e alle monadi leibniziane. E‘ interessante che nelle maggior parte dei testi l‘ etimologia di ergodo viene attribuita a ἔργον (energia) e ὁδός(via); mentre l‘ etimologia corretta è da ἔργον e

εἶδος, (perchè abbreviazione di ergomonodo, v. Ergodicità e Irreversibilità in questa collezione).

Moti Caotici: moti deterministici con proprietà caotiche, (v. Caos).

Moti Ordinati: moti regolari, periodici o quasi periodici (v. Teoria delle perturbazioni, Caos).

Moti Quasi Periodici: un moto è quasi periodico con pulsazioni fondamentali $\omega_1, \ldots, \omega_n$ se una qualunque funzione delle coordinate che lo descrivono, osservata in funzione del tempo sul dato moto, ha la forma $A(t) = f(\omega_1 t, \omega_2 t, \ldots, \omega_n t)$ ove $f(\varphi_1, \ldots, \varphi_n)$ è una funzione di n angoli, periodica in ciascuno con periodo 2π. Si suppone sempre che le pulsazioni siano *razionalmente indipendenti* (ossia che $\sum n_i\omega_i = 0$ sia possibile solo se $n_i = 0$ per ogni i). Le quantità $\nu_i = \omega_i/2\pi$ sono le frequenze fondamentali del moto quasi periodico: però le frequenze (o le pulsazioni) fondamentali di un MQP non sono univocamente definite. Se M è una matrice arbitraria a elementi interi e determinante ± 1 e se si pone $\omega'_i = \sum_j M_{ij}\omega_j$ si ottiene un altro possibile insieme di pulsazioni fondamentali. Una notevole proprietà delle pulsazioni findamentali ω_i è la densità dell' insieme dei punti aventi la forma $\sum n_i\omega_i$ al variare degli interi n_i. La trasformata di Fourier di $A(t)$ è una combinazione lineare discreta di funzioni delta di Dirac centrate sui punti ω aventi forma $\omega = \sum_i n_i\omega_i$ con n_i interi e i coefficienti di questa combinazione lineare sono legati semplicemente a quelli della trasformata di Fourier della f e sono in massima parte piccolissimi. Pertanto misure della trasformata di Fourier di un' osservabile, osservata su un moto quasi periodico, appaiono come una famiglia di picchi che emergono sul rumore di fondo; in numero sempre maggiore, e sempre più densi, al crescere della precisione delle misure; ma le posizioni dei picchi sono sempre esprimibili in termini delle pulsazioni fondamentali. Ovviamente se $n = 1$ il moto è periodico.

Moto Stazionario: è un moto in cui tutte le coordinate, e quindi tutte le grandezze osservabili variano in funzione del tempo in modo da avere un valore medio definito, indipendente dall' istante in cui se ne inizia la misura. Ad esempio, un moto periodico, è stazionario; ma anche un moto turbolento di un fluido è di solito un moto stazionario; se si osserva una grandezza in un punto qualsiasi si vede sì che cambia al variare del tempo, e senza mai riacquistare lo stesso valore, però evolve in modo da avere un valore medio ben definito. Uno stato stazionario dunque non va confuso con il suo caso particolare che è uno *stato di equilibrio*, nel quale nessuna grandezza fisica che descrive il sistema varia con il tempo. Nei sistemi dissipativi (come i fluidi), confinati a muoversi in un volume finito e soggetti a forze costanti o periodiche, si raggiunge uno stato stazionario dopo un transiente iniziale. Ma anche sistemi non limitati (come il mare, o un fiume) possono trovarsi in stati schematizzabili come stazionari, con ottima approssimazione, su

scale di tempo lunghe rispetto ai tempi caratteristici delle osservazioni, v. Caos, Turbolenza.

Multifrattalità: misura della variabilità della azione di espansione dei segmenti infinitesimi sotto l'azione delle iterate di una trasformazione S, regolare (differenziabile a tratti e localmente invertibile) di R^n in se. La S si suppone definita nell' intorno di un insieme A chiuso limitato e S–invariante (ossia $S\,A \subset A$). E l' azione di S è osservata su dati iniziali y scelti a caso rispetto ad una distribuzione di probabilità invariante μ definita su A e attribuente probabilità nulla agli (eventuali) insiemi di punti in cui S non è differenziabile (questo affinchè abbia sense definire l' azione di S sui segmenti infinitesimi). Dunque la multifrattalità è una proprietà di un sistema dinamico metrico (v.) definito da S, A, μ, ossia della terna (S, A, μ). Se $F_{1,k}(y)$ è il massimo coeffieciente di espansione, in y, delle lunghezze, rispetto alla trasformazioni S^k, iterata k-ma di S, si definisce la funzione del parametro reale α:

$$z(\alpha) = \lim_{k\to\infty} \frac{1}{k} \log \langle F_{1,k}(y)^\alpha \rangle$$

ove la media è su y ed è eseguita rispetto alla distribuzione di probabilità μ, se il limite esiste. La funzione $z(\alpha)$ è una misura della variabilità della espansione massima di S e contiene molta più informazione dell'esponente massimo di Lyapunov, che è dato da $z'(0)$ (se la derivata esiste). Ma questa non è l' unica definizione possibile di misura della multifrattalità di (S, A, μ). Ad esempio se μ è ergodica e m è il numero (indipendente da y) di esponenti di Lyapunov positivi e $F_{+,k}(y)$ è il massimo coefficiente di espansione degli elementi di volume di dimensione m per azione di S^k si pone:

$$\zeta(\alpha) = \lim_{k\to\infty} \frac{1}{k} \log \langle F_{+,k}(y)^\alpha \rangle$$

e questa è una altra misura della multifrattalità (ora $\zeta'(0) = \sum_{\lambda_j>0} \lambda_j$). Il sistema dinamico metrico (S, A, μ) si dice *multifrattale* se la funzione $z(\alpha)$ non è proporzionale ad α. Ad esempio se S è la trasformazione di $A = [0, 1]$ definita da $x \to 3x \bmod 1$ e $\mu(dx) = dx$ si trova che (S, A, μ) non è multifrattale e $z(\alpha) = \alpha \log 3$, mentre l'esponente di Lyapunov è $\lambda = \log 3$. Se S è la trasformazione di $A = [0, 1]$ definita da $x \to 3x \bmod 1$ per $x \in [0, \frac{1}{3}]$ e $x \to \frac{3}{2}(x - \frac{1}{3})$ per $x \in (\frac{1}{3}, 1]$ e $\mu(dx) = dx$ si trova che (μ è S–invariante e) (S, A, μ) è multifrattale e $z(\alpha) = \alpha \log 3 - \log \frac{1}{3}(1 + \frac{2}{2^\alpha})$ mentre l'esponente di Lyapunov è $\lambda = \frac{1}{3} \log 3 + \frac{2}{3} \log \frac{3}{2}$. La dipendenza da μ delle nozioni di esponente di Lyapunov e di multifrattalità è esemplificata dal seguente modello. Se S è la trasformazione di $A = [0, 1]$ definita da $x \to 3x \bmod 1$ per $x \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ e $x \to 27(x - \frac{1}{3}) \bmod 1$ per $x \in (\frac{1}{3}, \frac{2}{3})$ e $\mu(dx) = dx$ si trova che (μ è S–invariante e) (S, A, μ) è multifrattale e

$z(\alpha) = \log(\frac{2}{3}3^{\alpha} + \frac{1}{3}27^{\alpha})$ e l'esponente di Lyapunov è $\lambda = \frac{2}{3}\log 3 + \frac{1}{3}\log 27$. Ma si consideri il sistema dinamico (S, A, μ_1) con μ_1 che attribuisce misura 1 all' insieme dei numeri in $[0, 1]$ che sviluppati in base 3 non hanno la cifra 1 e che attribuisce misura 2^{-n} all' insieme dei numeri che non hanno la cifra 1 nel loro sviluppo in base 3 e che hanno le prime n cifre prefissate. Allora μ_1 è S–invariante e $z(\alpha) = \alpha \log 3$, e $\lambda = \log 3$: cioè (S, A, μ) è multifrattale e (S, A, μ_1) no.

Nernst , principio di W. Nernst: detto anche *terzo principio della termodinamica* dice che in un sistema condensato tutti gli stati di equilibrio termodinamico a temperatura nulla, che possono essere trasformati gli uni negli altri via trasformazioni isoterme reversibili, hanno entropia finita e uguale. Il problema è cosa si intenda per sistema condensato: è infatti ovvio che un gas perfetto classico con calore specifico $C_V = 3R/2$ non verifica questo principio, (perchè la sua entropia per particella diverge per $t \to 0$ e, comunque, dipende esplicitamente dalla densità ρ: $s = \frac{3}{2}R\log T - R\log\rho$), come non lo verifica il sistema ideale classico costituito da atomi che interagiscono via un'interazione di puro cuore duro (essenzialmente per gli stessi motivi del gas perfetto classico). Il principio di Nernst non è un teorema nella meccanica statistica, e si ritiene che debba valere senza eccezioni in sistemi per cui valga la meccanica quantica. Come si è detto, nella meccanica statistica classica ci sono ovvie eccezioni: altre eccezioni si verificano in sistemi che sono almeno parzialmente modellabili seguendo seguendo la meccanica classica. Ma se un sistema viene pensato come descritto da ioni ed elettroni completamente quantizzati (e non descritti in modo semiclassico) allora il principio deve essere valido senza eccezioni. Matematicamente è connesso alla proprietà dell' equazione di Schrödinger di un sistema con un arbitrario numero di particelle descritto da un potenziale di interazione stabile (v.), di avere uno stato fondamentale non degenere, per tali sistemi l' entropia microcanonica sarebbe nulla e quindi il principio sarebbe valido. Ma questa proprietà è vera solo per sistemi di particelle verificanti, se identiche, la statistica di Bose e senza stuttura (cioè senza spin): poichè i sistemi che si incontrano in natura sono invece costituiti da fermioni con spin, si vede che il principio non è in realtà un teorema di meccanica statistica, ma fornisce informazioni ulteriori e per questo è stato a volte posto in dicussione. Una sua conseguenze è che il calore specifico C_V si annulla per $T \to 0$ in modo che $C_V dT/T$ sia integrabile (altrimenti l' entropia divergerebbe). Il principio permette di assegnare entropia nulla allo stato di vuoto e quindi una entropia assoluta a tutti gli stati del sistema. A temperatura nulla l' entropia di tutti gli stati connessi da trasformazioni reversibili ha un valore che è chiamato l' entropia residua degli stati in questione. Ad esempio allo zero assoluto il sistema composto da idrogeno, ossigeno e

acqua ha, nella fase acqua, una entropia residua non nulla (e il suo valore è spiegato dal modello del ghiaccio di Pauling (1935), v.). Il principio di Nernst è a volte chiamato *teorema del calore* di Nernst: che lo dedusse dal secondo principio e dall' ipotesi che i calori specifici (ad esempio C_V nel caso di un sistema con parametri di stato P, V, o C_H nel caso di una sostanza paramagnetica) dei corpi condensati tendano a zero per $T \to 0$, ovvero dall' ipotesi che non possa esistere un trasformazione adiabatica che produca il raffreddamento del corpo da temperatura positiva a temperatura nulla.

Numeri Primi, teorema dei: stabilisce che il numero $\mathcal{N}_N$ di numeri primi $\leq N$ è, asintoticamente per $N \to \infty$, dato da $\mathcal{N}_N \simeq \mathcal{N}_N^0 = \frac{N}{\log N}$, nel senso che il rapporto dei due membri tende a 1. In particolare la densità dei numeri primi contenuti in $[1, N]$ tende a zero per $N \to \infty$. Una forma più primitiva è il teorema di Eulero che afferma che la serie $\sum_{p=primo} \frac{1}{p} = +\infty$. Una forma più precisa è il teorema di Hadamard e de la Vallèe Poussin che afferma che $\delta(N) = |\mathcal{N}_N - Li(N)|/Li(N)| \leq O(e^{-c\sqrt{\log N}})$ con $c > 0$ oppurtuna e $Li(N)$ è la funzione logaritmo integrale $Li(N) = \int_0^N \frac{dt}{\log t}$. Questi teoremi si possono dimostrare in base alle proprietà di analiticità della funzione zeta di Riemann. La stessa ipotesi di Riemann (assenza di zeri non banali, della funzione zeta di Riemann, con parte reale $> 1/2$) è *equivalente* alla affermazione che $\delta(N) \leq O(\frac{1}{\sqrt{N}}(\log N)^2)$ (o alla affermazione apparentemente più debole che $\delta(N) \leq O(\frac{1}{N^{1/2-\varepsilon}})$ per ogni $\varepsilon > 0$). *Bibliografia*: H. Edwards, *Riemann's zeta function*, Academic Press, New York, 1974.

Olodo: termine usato da Boltzmann per denotare quello che oggi è un elemento dell' insieme statistico canonico (v. Distribuzione canonica, IS).

Ornstein, teorema di D. Ornstein: due schemi di Bernoulli con uguale entropia sono isomorfi. Esiste cioè un codice che trasforma le successioni campione dell' uno in quelle dell' altro in modo che la statistica di Bernoulli delle prime sia trasformata in quella delle seconde: questa proprietà è vera se si escludono dalla corrispondenza stabilita dal codice opportuni insiemi di successione campione che però hanno probabilità nulla nei rispettivi schemi di Bernoulli, (v. Entropia e Informazione).

Ornstein–Zernike, comportamento di L. Ornstein, F. Zernike: in un gas classico in equilibrio termodinamico a temperatura superiore alla temperatura critica e con interazione intermolecolare a corta portata la funzione di correlazione spaziale troncata, a due punti, $h(r)$ (v. funzione di Correlazione spaziale) si comporta, per $r \to \infty$ come $r^{-(d-2)}e^{-\kappa r}$ ove d è la dimensione spaziale e κ^{-1} è la lunghezza di correlazione, v. Esponenti critici.

Ortodo: termine usato da Boltzmann per denotare un insieme statistico i cui elementi forniscono un modello microscopico della Termodinamica

classica macroscopica, (v. MSC, IS).

Osservabile Locale: grandezza osservabile in un sistema di particelle che occupa un volume V, finito o infinito, e che dipende solo dalle coordinate delle particelle contenute in un volume $\Lambda \subset V$ finito: tale grandezza si dice localizzata in Λ. Ad esempio il numero di particelle in una piccola regione diviso per il volume della regione stessa definsce la densità locale (da non confondersi concettualmente con la densità totale del sistema, N/V), v. MSC, IS, Limite Termodinamico.

Paradossi, della MSC: v. MSC, §8 e segg.

Partizione, funzione di: v. Funzione di P..

Peierls, argomento di Peierls: dimostrazione della esistenza di transizioni di fase nel modello di Ising (v.) a dimensione spaziale $d \geq 2$. La tecnica è però molto più generale ed ha condotto alla comprensione di un gran numero di transizioni di fase semplici con rottura di simmetria (v.), e anche ad una molto dettagliata analisi microscopica del fenomeno di coesistenza delle fasi e della associata tensione superficiale (nei casi cui si applica): v. Transizioni di Fase, Tensione Tuperficiale. L' appellativo di *argomento* è dovuto al fatto che la teoria di Peierls (1937) non risponde forse ai criteri di rigore della Fisica Matematica ed è stato posto in forma matematica pienamente rigorosa da R. Dobrushin (1965) acquisendo in questa forma il nome di metodo di Peierls.

Peso Statistico: è una quantità definta per ogni evento e proporzionale alla probabilità dell' evento stesso. E' utile per definire una distribuzione di probabilità quando siano interessanti solo i rapporti fra probabilità di eventi. In ogni caso dividendo il peso statistico per la somma dei pesi statistici di tutte le configurazioni si ottiene la probabilità. La somma dei pesi stastistici si chiama funzione di partizione: evidentemente non è univocamente definita dalla distribuzione di probabilità. Tuttavia in molti casi si fissano i pesi statistici secondo convenzioni ben definite e allora la funzione di partizione può avere un significato interessante. Questo avviene nei casi degli insiemi statistici della MSC ove il peso statistico delle configurazioni è fissato uguale al fattore di Boltzmann (v. Fattore di Boltzmann, Distribuzione microcanonica, canonica, gran canonica).

Pesin, formula di P.: relazione, valida in vari casi, fra entropia di un sistema dinamico metrico (v.) ed esponenti di Lyapunov. (v.); v. Termalizzazione.

Poincaré, cicli di H. Poincarè: v. Ricorrenza.

Potenziale, stabile: nella MSC un potenziale di interazione (v. Interazione) si dice stabile se l' energia potenziale di una configurazione qualsiasi con n particelle $V = V(\underline{r}_1, \ldots, \underline{r}_n)$ verifica la disuguaglianza $V \geq -Bn$ ove B è una costante indipendente dalla configurazione e da n. Un esempio di P.S. è ovviamente un potenziale non negativo; un esempio meno banale

è un potenziale a due corpi (v. Interazione) che si possa scrivere come somma di un potenziale non negativo (*repulsivo*) e di un potenziale con trasformata di Fourier non negativa (Fisher, Ruelle). Un potenziale a due corpi continuo e negativo nell' origine è invece instabile perchè ovviamente la configurazione in cui tutte le particelle sono poste nello stesso punto ha energia negativa dell' ordine del quadrato del numero di particelle. Il potenziale gravitazionale (o quello di Coulomb) non è stabile in MSC. La nozione di stabilità non richiede che le n particelle siano identiche. Nella MSQ la nozione corrispondente è che il livello fondamentale dell' operatore di Hamilton per n particelle (o l' estremo inferiore dello spettro) sia $\geq -Bn$. La nozione quantica è profondamente diversa da quella classica perchè nell' operatore hamiltoniano è presente l' energia cinetica che da un contributo importante sia a causa del principio di indeterminazione sia a causa delle proprietà di simmetria delle funzioni d' onda dipendenti dalla natura fermionica o bosonica delle particelle (mentre classicamente la proprietà di stabilità non muterebbe natura se si includesse nell' energia anche l' energia cinetica, perchè quest' ultima dovrebbe essere posta uguale a zero per il calcolo dell' energia minima). L' importanza della nozione di stabilità sta nel fatto che ci si può aspettare che un sistema macroscopico si comporti secondo quanto familiare dalla Termodinamica macroscopica solo se l' interazione microscopica è stabile (v. Stabilità della materia). Di grande importanza è quindi il teorema di F. Dyson–A. Lenard secondo il quale un sistema di un numero finito di specie di particelle cariche quantiche, costituito da fermioni di cariche arbitrarie (e, eventualmente, bosoni purchè con cariche tutte dello stesso segno), è stabile (v. Stabilità della materia). In meccanica statistica classica, invece, un sistema di un numero finito di specie di particelle cariche è stabile solo se in aggiunta all' interazione coulombiana è presente una interazione a cuore duro fra tutte le specie di particelle (L. Onsager).

Potenziale, superstabile: in MSC un potenziale a due corpi si dice superstabile se l' energia potenziale di una configurazione qualsiasi con n particelle $V = V(\underline{r}_1, \ldots, \underline{r}_n)$ contenute in un volume Λ verifica la disuguaglianza $V \geq -Bn + An^2/\Lambda$ ove A, B sono costanti indipendenti dalla configurazione e da n; di solito si richiede che la disuguaglianza sia valida per regioni Λ che non sono più piccole di un cubetto Λ_0, fissato una volta per tutte. Un esempio di P. stabile ma non superstabile è ovviamente il potenziale identicamente nullo; ma la gran parte dei potenziali a due corpi e a corta portata, di rilevanza per la fisica, è superstabile (in particolare lo sono i potenziali di Lennard Jones). Una delle proprietà più semplici dei potenziali superstabili è la dipendenza continua della pressione dalla densità (R. Dobrushin–R. Minlos). I potenziali a cuore duro sono ovviamente superstabili, se stabili (in particolare se a corta portata). La definizione di

superstabilità in MSQ viene data in modo simile a quella della stabilità richiedendo che l' estremo inferiore dello spettro dell' hamiltoniana del sistema racchiuso in una regione Λ sia non inferiore a $-Bn+An^2/\Lambda$: quì però occorre specificare anche le condizioni al contorno che si usano per definire l' operatore energia cinetica e restringere le forme permesse a Λ e non c' è una prescrizione generale perchè la teoria non è stata molto sviluppata.

Principio Variazionale, in MSC: lo stato (o gli stati) di equilibrio ter modinamico con temperatura T e potenziale chimico μ verifica, sotto opportune ipotesi restrittive sull' interazione, la proprietà di massimizzare fra tutti gli stati γ possibili invarianti per traslazione (ossia fra tutte le possibili distribuzioni di probabilità invarianti per traslazione definite sullo spazio delle fasi) la differenza fra entropia per unità di volume $s(\gamma)$ e $\beta(u(\gamma) - \mu\rho(\gamma))$, ove $\beta = 1/k_BT$, u è l' energia per unità di volume e ρ la densità. La dimostrazione della validità di questa affermazione discende dall' osservazione (di semplice verifica) della sua validità nelle distribuzioni gran canoniche di sistemi confinati in un volume finito. v. Limiti termodinamici, MSC.

Principio Variazionale di Ruelle: v. Caos.

Processi puntuali: distribuzioni di pribabilità su insiemi di eventi ognuno dei quali consiste in una famiglia (numerabile o finita) di punti in uno spazio R^d (o più in generale in una varietà a d dimensioni). Ad esempio gli stati di Gibbs della MSC dei sistemi monoatomici si possono considerare come processi puntuali sullo spazio R^6 delle posizioni e impulsi delle particelle, v. Limiti Termodinamici, Processi Stocastici.

Punto Critico, di un sistema in equilibrio termodinamico; punto nello spazio degli stati di equilibrio termodinamico nell' intorno del quale avvengono fenomeni critici (v.), quando i parametri di stato vengono fatti tendere ai "valori critici". A secondo del modo in cui, nello spazio degli stati di equilibrio termodinamico, ci si avvicina al punto critico si osservano fenomeni critici diversi, (v.). Ad esempio se ci si avvicina al punto critico a volume costante si osserva di solito, come fenomeno critico, che il calore specifico diverge; se ci si avvicina al punto critico a temperatura costante si osserva che la comprimibilità diverge. In generale si tende a descrivere i fenomeni critici in termini di divergenza di grandezze termodinamiche. Le grandezze termodinamiche che divergono al punto critico sono spesso grandezze che misurano le fluttuazioni di altre grandezze termodinamiche. Ad esempio la divergenza della comprimibilità significa che al punto critico le fluttuazioni di densità sono più grandi del loro valore *normale*: in punti non critici il numero di particelle in un volume V è proporzionale a V con fluttuazioni dell' ordine di $V^{1/2}$ mentre al punto critico è ancora proporzionale a V ma con fluttuazioni proporzionali a $V^{1/2}V^{(2-\eta)/2d}$, ove η è un esponente critico (v.) e d la dimensione dello spazio ($d = 3$, di solito), (v. Fenomeni Critici,

Esponenti Critici).

Quadratura: integrale definito unidimensionale. La locuzione *riduzione alle quadrature* di un problema significa che la soluzione del problema è ridotta la calcolo di un numero finito (piccolo) di integrali definiti di funzioni che si possono considerare ben note. Esempi di problemi risolti da quadrature sono la determinazione dei moti dei sistemi integrabili elementari della meccanica classica o la soluzione dei modelli risolubili della MS, v. Teoria delle Perturbazioni, Meccanica Classica, Modelli Risolubili.

Radiazione nera: v. Corpo nero, Statistiche Quantiche.

Rayleigh–Jeans, formula di R.J.: dà la densità di energia presente per unità di intervallo di frequenza nella radiazione di un corpo nero, nell' ipotesi che valga la equidistribuzione dell' energia. Tale densità è: $u_\nu = 8\pi\nu^2 k_B T/c^3$, (ove k_B, c, T sono rispettivamente la costante di Boltzmann, la velocità della luce e la temperatura assoluta), v. MCS, Critica della Meccanica Statistica, Statistiche Quantiche, Legge di Planck.

Reversibilità: v. Irreversibilità, MSC, Legge di Planck, Demonietto di Maxwell.

Ricorrenza, tempo di R.: in un sistema di N particelle chiuso in un contenitore a pareti idealmente riflettenti e che evolve secondo le equazioni di Hamilton, dato un dato iniziale non di equilibrio x e una precisione ε comunque piccola, esiste un dato iniziale vicino entro ε a x, che si allontana da x più di ε e poi, dopo un tempo t_ε abbastanza lungo, ritorna vicino a x entro ε. Questo è il teorema di ricorrenza di H. Poincarè e t_ε è *un tempo di ricorrenza* con precisione ε. Se lo spazio delle fasi viene sostituito con uno spazio discreto (v. Celletta, Dinamica Discretizzata) e se si può assumere valida l' ipotesi ergodica (v. Ipotesi Ergodica di Boltzmann) è possibile stimare il tempo di ricorrenza relativo alla scelta della precisione ε della ricorrenza, con ε= dimensione delle cellette. Tale tempo risulta già in sistemi con poche decine di particelle e in sistemi che modellano sistemi di interesse nella teoria dei gas, dell' ordine di grandezza di innumerevoli età dell' universo. Questa osservazione di Boltzmann toglie ogni possibilità di utilizzare il teorema di R. per mostrare la impossibilità dell' irreversibilità in sistemi a dinamica microscopica reversibilie (v. Irreversibilità, Demonietto di Maxwell, MSC §3,(4.3)).

Rottura Spontanea di Simmetria, nelle transizioni di fase: in generale si considera una equazione $F(x) = 0$ per un oggetto x e si suppone che sullo spazio in cui x può variare agisca un gruppo G di trasformazioni. Denotando gx il punto in cui x è trasformato dall' elemento $g \in G$ si suppone che $F(gx) \equiv F(x)$ per ogni $g \in G$. Si dice allora che l' equazione $F(x) = 0$ è invariante rispetto al gruppo G. Ci si aspetta che le soluzioni x siano tali che $gx = x$ per ogni g. Tuttavia può accadere che così non sia e allora se x è una tale soluzione si dice che la simmetria dell' equazione

è spontaneamente rotta; ovviamente in questi casi esistono altre soluzioni (ottenibili da x agendo su x con gli elementi del gruppo G). Un tipico esempio importante di equazioni che possono presentare il fenomeno della rottura spontanea di simmetria è fornito dalle equazioni DLR, BBGKY, KS, KMS (v.) e altre le cui soluzioni determinano le funzioni di correlazione degli stati di Gibbs, ossia di equilibrio termodinamico, in sistemi infiniti (occupanti cioè l' intero spazio), (v. Limite termodinamico). Le soluzioni di queste equazioni, in tali casi, sono interpretabili come fasi diverse del sistema e quindi la rottura spontanea della simmetria fornisce un mezzo per riconoscere l' esistenza di una transizione di fase. Le transizioni di fase più semplici sono esempi di rottura spontanea della simmetria (ad esempio, nel ferromagnetismo, ove si trovano, in campo magnetico esterno nullo, stati di equilibrio con magnetizzazione media non nulla). Però esiste la possibilità che $F(x) = 0$ abbia più di una soluzione e che tuttavia $gx = x$ per ogni soluzione. Nel caso delle equazioni di equilibrio termodinamico di sistemi infiniti questo corrisponde a transizioni di fase senza rottura spontanea della simmetria. Esempi di sistemi con transizioni di fase dovute a rottura spontanea di simmetria sono il modello di Ising in campo nullo a dimensione $d \geq 2$ (v.), il modello di Widom–Rowlinson (v.), il modello di Heisenberg (v.) antiferromagnetico a $d \geq 3$, e (si crede) il modello di Heisenberg ferromagnetico a $d \geq 3$. Mentre la transizione liquido–gas in un gas omogeneo dovrebbe essere un esempio di transizione senza rottura spontanea di simmetria (perchè non ha nessuna simmetria), come pure potrebbe essere tale il modello di Heisenberg a $d = 2$ (che pur avendo una simmetria non la può rompere spontaneamente, per il teorema di Mermin–Wagner, e che si congettura presenti una transizione di fase di altro tipo).

Scala, leggi di: v. Leggi di scala, Esponenti critici.

Scenarii, per lo sviluppo del caos: v. Caos.

Seme, di una successione casuale: dato iniziale $x \in [0, 1]$ per una trasformazione S tale che $S^n x$ è una successione di punti in $[0, 1]$ distribuiti a caso con distribuzione uniforme su $[0, 1]$, (v. Metodo Montecarlo, Generatori di numeri casuali, Caos).

Sfere dure, modello di gas: v. cuore duro.

Shannon,teorema di S.: v. Entropia e Informazione.

Simplesso , insieme convesso I di R^n tale che ogni punto può essere ottenuto come baricentro di un' unica distribuzione di masse sui punti estremali di I. Questa nozione si estende a spazi vettoriali topologici a dimensione infinita, purchè con topologia non troppo strana, e si incontra nella teoria delle fasi pure; in particolare l' insieme degli stati di Gibbs associati ad una data interazione in un modello di Ising (v.) forma un simplesso in un senso naturale e questo ha l' interpretazione fisica che gli stati di equilibrio termodinamico sono pensabili in un unico modo come

miscugli di fasi pure, v. Limite Termodinamico.

Simulazione numerica metodi, v. Termalizzazione, Dinamica Discretizzata, Coefficienti di Trasporto.

Sinai, Biliardo di: un punto materiale chiuso in un quadrato a pareti perfettamente riflettenti e contenente uno o più ostacoli circolari (o almeno strettamente convessi). L' insieme $\underline{\omega}$ dei dati iniziali con velocità di modulo 1 è considerato come lo spazio delle fasi di un sistema dinamico metrico in cui la evoluzione temporale è semplicemente il moto rettilineo uniforme del punto eccetto che agli istanti di collisione con le pareti o con gli ostacoli quando si suppone che abbia luogo un urto elastico perfetto. Se ϑ è l' angolo che la velocità forma con un asse fisso e se (x, y) sono le coordinate di posizione e se V è la superficie del quadrato privato degli ostacoli, un punto generico dello spazio delle fasi è individuato da $\xi = (x, y, \vartheta)$ e l' evoluzione conserva il volume $\mu(dx) = dx\, dy\, d\vartheta/(2\pi V)$. Se dunque $S_t\xi$ denota il punto in cui ξ evolve nel tempo $t > 0$ il sistema dinamico $(\underline{\omega}, \mu, S_t)$ è un sistema dinamico metrico. La sua importanza sta nell' essere il primo (e sostanzialmente a tutt' oggi, l' unico) sistema dinamico direttamente dotato di interpretazione fisica del quale è stato possibile dimostrare l' ergodicità (Sinai, 1961). Si può dimostrare di più, ad esempio che il sistema è un generatore di numeri casuali (v.) nel senso che segue. Si divide lo spazio delle fasi in n, ad esempio $n = 2$, insiemi $I_1, \ldots, I_n$ di misura positiva e si assegna un dato iniziale a caso, con distribuzione (ad esempio) μ cioè uniforme. Fissata un' unità di tempo t si genera la successione $\underline{\sigma}$, storia di x, ponendo $S_{it}x \in P_{\sigma_i}$ per $i = 0, 1, \ldots$. Allora la successione $\underline{\sigma}$ ha, con probabilità 1, frequenze definite (v.), entropia (v.) positiva, ed è codificabile in uno schema di Bernoulli (v. Entropia e Informazione, Bernoulli).

Stabilità della Materia: v.

Stabilità, hamiltoniana in MS: v. Potenziale stabile e superstabile, Stabilità della materia.

Stati, puri e misti in MS: v. Limite Termodinamico, Transizioni di Fase.

Statistica di una successione di simboli: si consideri una successione $\underline{\sigma} = (\sigma_i)_{i=0,1,\ldots}$ di simboli $a_1, a_2, \ldots, a_p$ presi da un insieme (alfabeto) A finito e si supponga che $\underline{\sigma}$ abbia frequenze definite (v. Frequenza di una Stringa). La statistica della successione è la collezione delle frequenze di tutte le stringhe finite. La statistica di una successione a frequenze definite può convenientemente essere pensata come una misura di probabilità sullo spazio prodotto $B = A^{Z_+}$, cioè sullo spazio di tutte le possibili successioni formate con elementi dell' alfabeto A. La definizione sta semplicemente nel dire che gli insiemi misurabili sono generati dagli insiemi di tutte le successioni i cui primi n simboli coincidono con una prefissata stringa $\underline{a}$ di n simboli (al variare du n e della stringa $\underline{a} = (a_0, \ldots, a_{n-1})$ di n simboli di A). Denotiamo $C_{\underline{a}}$ questi insiemi; e la misura, ossia la probabilità, di

questi insiemi $C_{\underline{a}}$ è per definizione la frequenza della stringa $\underline{a}$ in $\underline{\sigma}$. Si può verificare che la distribuzione μ così definita su B è un processo stocastico, (v. Entropia e Informazione, Processi Stocastici).

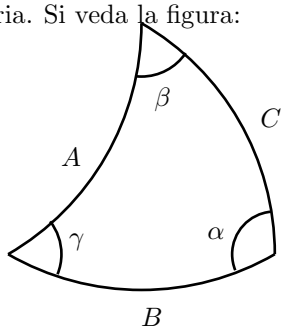
Stefan–Boltzmann, costante di S.B.: nella teoria del corpo nero è il rapporto $\sigma = S/T^4$ fra la energia S emessa in un semispazio (cioè su un angolo solido 2π) nell' unità di tempo da una unità di superficie di corpo nero in equilibrio termodinamico a temperatura assoluta T e la quantità T^4 (legge di Stefan–Boltzmann). L' esistenza della costante è conseguenza delle leggi dell' elettromagnetismo e dei principi della termodinamica; il valore della costante dipende però dalla legge secondo la quale la energia si distibuisce nella radiazione alle varie frequenze. Se si suppone valida la legge di Planck si trova:

$$\sigma = \frac{2\pi 5 k_B^4}{15 c^2 h^3}$$

ove k_B, c, h sono, rispettivamente, la costante di Boltzmann, la velocità della luce e la costante di Planck.

Superfici Aleatorie: analogo bidimensionale dei cammini aleatorii (v.). Esempi tipici di superfici aleatorie si incontrano nella teoria microscopica della coesistenza delle fasi; più recentemente hanno acquisito molta importanza nella teoria delle Stringhe nella Fisica quantistica relativistica, (v. Tensione superficiale, Stringhe)

Trigonometria sferica: è una generalizzazione della trigonometria piana intesa come teoria metrica dei triangolo. Sulla sfera unitaria si definisce *retta* un cerchio massimo, *segmento* un arco di cerchio massimo. Due cerchi massimi hanno in comune, sempre, 2 punti distinti, a meno che non coincidano. Due cerchi massimi orientati formano, nei punti di intersezione, un angolo che è misurato dall' angolo piano formato dalle tangenti in uno dei punti di intersezione. Tre archi orientati i cui estremi sono due a due coincidenti, in modo che ci siano solo tre punti comuni ai tre segmenti, e che inoltre hanno orientazioni concordi, formano un *triangolo sferico*, che è definito come la parte di superficie della sfera che vede l' orientazione dei tre archi come antioraria. Si veda la figura:



La trigonometria sferica si basa sulle seguenti identità fondamentali:

$$\frac{\sin A}{\sin \alpha} = \frac{\sin B}{\sin \beta} = \frac{\sin C}{\sin \gamma} \qquad \text{teorema dei seni}$$

$$\cos A = \cos B \cos C + \sin B \sin C \cos \alpha \qquad \text{teorema del coseno}$$

$$\cos A \cos \gamma = \sin A \cot B - \sin \gamma \cot \beta \qquad \text{teorema delle 4 parti}$$

$$\sin C \cos \beta = \cos B \sin A - \sin B \cos A \cos \gamma \qquad \text{teorema analogo}$$

e ovviamente se si operano le sostituzioni $(A, \alpha) \to (B, \beta) \to (C, \gamma) \to (A, \alpha)$ le relazioni che si ottengono dalle precedenti sono pure vere, e così pure le relazioni in cui due delle tre coppie sono pernutate fra loro (simmetrie rotazionale e antipodale).
La trigonometria sferica si può estendere a superfici generali in cui il ruolo dei segmenti è preso dagli archi di geodetica e le rette sono le geodetiche: lo studio delle relazioni fra geodetiche, segmenti di geodetica, ecc., costituisce un semplice esempio di geometria non euclidea (v. Geometie non Euclidee).
Il *teorema di Gauss* dice che, nel caso della trigonometria sferica, l' area di di un triangolo Δ è legata al *difetto geodetico*:

$$\alpha + \beta + \gamma = \pi + C \operatorname{area}(\Delta)$$

ove C è una costante. In generale questa formula è valida per triangoli su superfici arbitrarie con $C \operatorname{area}(\Delta)$ sostituito da $\int_\Delta C(x) d\sigma(x)$, ove $C(x)$ è una funzione detta *curvatura gaussiana* e $d\sigma(x)$ è l' elemento di superficie. Nel caso della sfera unitaria la funzione $C(x)$ è costante per ragioni di simmetria e può essere calcolata considerando il triangolo sferico ottenuto marcando su un cerchio massimo orientato tre punti: si vede che in questo particolare triangolo la somma dei tre angoli al vertice è 3π mentre l' area è 2π: quindi $C = 1$. Se invece della sfera unitaria si usa una sfera di raggio R si trova $C = R^{-2}$, e in generale C si dice *curvatura gaussiana*. E si verifica facilmente che la trigonometria sulla sfera di raggio R diventa, nel limite $R \to \infty$, la familiare trigonometria piana.
La trigonometria sferica gioca un ruolo importante nella Astronomia, nella Geodesia e in generale nella Meccanica Analitica (v. Sistemi Rigidi).

Universalità, al punto critico: v. Fenonmeni critici.

Universalità, nei moti caotici: v. Caos.

van der Waals, equazione di stato di J. van der Waals: v. Serie del Viriale, MSC, Potenziale di Kac, Transizioni di Fase, Esponenti Critici.

Variabili a blocco: v. Fenomeni Critici.

Vincolo Ideale: un sistema di n punti materiali si dice soggetto ad un vincolo ideale olonomo (etimologicamente "legge globale" da ὅλος(globale) e νόμος(legge)) se:

(1) i suoi punti interagiscono in modo da muoversi verificando sempre un certo numero di relazioni, $\varphi_1(\underline{x}_1, \ldots, \underline{x}_n) = 0, \ldots, \varphi_p(\underline{x}_1, \ldots, \underline{x}_n) = 0$, dette *equazioni dei vincoli*, qualunque siano le altre forze, *forze esterne*, che vengono applicate al sistema, purchè tali relazioni siano verificate all' istante iniziale.
(2) se si applicano al sistema forze conservative di energia potenziale $V(\underline{x}_1, \ldots, \underline{x}_n)$ le soluzioni delle equazioni del moto che occupano le configurazioni $X^1 = (\underline{x}_1^1, \ldots, x_n^1)$ all' istante t_1 e $X^2 = (\underline{x}_1^2, \ldots, \underline{x}_n^2)$ all' istante t_2 sono le funzioni $t \to X(t)$ che rendono stazionaria l' azione:

$$\int_{t_1}^{t_2} \Big[T(\dot{X}(t)) - V(X(t)) \Big] \, dt$$

nell' insieme dei moti che si svolgono fra t_1 e t_2 con configurazioni estreme X^1 e X^2.
Poichè i vincoli sono di solito realizzati, nella realtà, a mezzo di forze elastiche molto intense, che realizzano il vincolo nel limite in cui la intensità diviene infinita, è necessario verificare caso per caso se le forze che realizzano il vincolo sono tali da permettere di chiamare il vincolo ideale, con il conseguente vantaggio della possibilità di dedurre le equazioni del moto dal principio variazionale di cui al punto (2).
In modo analogo si definiscono i "vincoli anolonomi", come vincoli in cui le relazioni φ_i coinvolgono anche le velocità (o gli impulsi): ad esempio se F è una quantità conservata, allora $F - f$ è un vincolo per ogni valore della ostante f, (cioè se $F - f = 0$ all' istante iniziale resta tale agli istanti successivi). Così le equazioni $\dot{\underline{x}}_j = \underline{p}_j/m$, $\dot{\underline{p}}_j = \underline{F}_j + \underline{E} - \alpha(\underline{p})\underline{p}_j$, $j = 1, \ldots, N$, con $\alpha(\underline{p})$ dato da $\underline{e} \cdot \sum_j \underline{p}_j) / \sum_j \underline{p}_j^2$ e $\underline{F}_j = \underline{\partial}_{\underline{x}_j} V(\underline{x}_1, \ldots, \underline{x}_N)$ è una forza conservativa, allora $H - f \equiv (\sum_j \underline{p}_j^2/2m + V) - f$ è un vincolo anolonomo, per ogni f, perchè si vede subito che H è una quantità conservata.

Viriale, di un sistema di forze: in un sistema di N particelle, che occupano le posizioni $\underline{r}_1, \ldots, \underline{r}_n$ e sono soggette a forze rispettive $\underline{f}_1, \ldots, \underline{f}_n$, il V. delle forze è la somma $-\sum_i \underline{x}_i \cdot \underline{f}_i$. Se la risultante delle forze è nulla allora il V. è indipendente dalla scelta dell' origine delle coordinate (il caso in cui le forze fra le varie particelle siano solo forze interne verificanti il terzo principio della dinamica è dunque uno di questi casi).

Viriale, teorema o equazione del V.: in un sistema meccanico a N particelle (non necessariamente identiche), interagenti con un potenziale a due o più corpi e confinato in un volume finito, il valore medio dell' energia cinetica è, su ogni moto del sistema, uguale alla metà del valore medio del viriale delle forze (Clausius). Questo teorema è stato il punto di partenza della teoria di van der Waals. Si supponga che l' interazione fra le particelle consista in una parte a cuore duro di raggio $a > 0$, in una parte $\varphi(\underline{r})$ a lun-

ga portata, attrattiva ma di debole intensità e in una parte che esprime le forze esterne che le pareti esercitano sulle particelle per mantenerle nel volume V di confinamento. Per lunga portata si intende qui che il numero di particelle del sistema contenute in una sfera di raggio uguale alla portata del potenziale è molto grande, di modo che, ai fini del calcolo della forza subita da una data particella non è mai importante la posizione di ogni singola altra particella. Si vede allora che il viriale medio è somma di tre termini dovuti alle tre forze menzionate. Si trova che, se P, V, ρ, T, k_B denotano la pressione, il volume, la temperatura, la costante di Boltzmann e se $v_0 = (4\pi a^3/3)$ è il volume delle molecole e $A = 2^{-1} \int \varphi(\underline{r}) d\underline{r}$ (*campo medio* cui è soggetta la particella generica), e $\beta = 1/k_B T$:
(1) il viriale medio delle forze delle pareti è $3PV$, (dalle definizioni di P e dalla formula di Green $\int_{\partial V} \underline{x} \cdot \underline{n}\, d\sigma \equiv 3V$)
(2) il viriale medio delle forze interne repulsive di cuore duro è $\frac{-3}{\beta}(4v_0\rho^2 V) + O(\rho^3)$, come si può vedere perchè per lo stesso motivo del punto precedente è $-3P8v_0\frac{N}{2}$ e $P = 1/\beta v$ a meno di correzioni di $o(1/v)$; il segno $-$ proviene dal fatto che le collisioni son "esterne" alle sferette e il fattore 8 è perchè il volume occupato da una particella impenetrabile di raggio a "rispetto alle altre" è quello di una sfera di raggio $2a$; infine $N/2$ appare in luogo di N per evitare di contare due volte una stessa collisione fra due particelle.
(3) il viriale medio delle forze attrattive di lunga portata è $3VA\rho^2 + O(\rho^3)$, (v. serie del viriale, potenziale di Kac, transizioni di fase). Si vede dunque che il calcolo del viriale, in prima approssimazione, conduce ad una correzione alla legge dei gas perfetti che coincide al secondo ordine nella densità con quella della equazione di van der Waals (v. Serie del viriale, formula (10)): le correzioni di ordine superiore alla equazione di stato possono essere calcolate e si ottiene cosi' la serie del V. (v.). La coincidenza con la equazione di van der Waals cessa però agli ordini superiori, tranne che se la dimensione dello spazio è $d = 1$ e il potenziale è il caso limite di un potenziale di Kac (v.). Se $d > 1$ la equazione di van der Vaals, pur non essndo strettamente corretta, coglie tutti gli aspetti della teoria (inclusi gli esponenti critici principali) *purchè* ancora il potenziale di interazione a lunga portata sia il caso limite di un potenziale di Kac (v.); altrimenti la teoria ha poco a che fare con l' equazione di stato dei gas reali almeno nell' intorno del punto critico, pur fornendo una non irragionevole descrizione dell' equazione di stato lontano dal punto critico in termini di due soli parametri; al punto che spesso una equazione di stato di un gas reale viene scritta nella forma di van der Waals pensando i parametri A, B come funzioni di T, ρ e tabulandone le (relativamente piccole) variazioni al variare di T e ρ. L' equazione di van der Vaals può essere usata per ottenere stime dei parametri atomici fondamentali (raggio delle molecole e intensità delle forze intermolecolari); v. MSC, serie del viriale, Potenziale di Kac, transi–

zioni di fase, numero di Avogadro).

Viscosità: dinamica, cinematica, relazioni di Clausius–Maxwell: la viscosità dinamica η è definita come il rapporto fra il gradiente della velocità e la forza orizzontale per unità di superficie necessaria a mantenere il gradiente di velocità stesso in un moto stazionario di un fluido infinito che scorra in modo da mantenere ovunque velocità parallela ad un dato asse. La viscosità cinematica si definisce come il rapporto η/ρ fra viscosità dinamica e densità. Il cammino libero medio λ è legato alla viscosità dinamica η dalle relazioni di Clausius–Maxwell): $\eta = \rho\lambda v/3$, $\lambda = m/(4\pi\rho a^2\sqrt{2})$ ove m è la massa delle particelle, ρ è la densità, v è la velocità media $v = (3k_BT/m)^{1/2}$ e a è il raggio delle molecole. Valori tipici della viscosità dinamica sono (a 20^oC) dell‘ ordine di $10^{-2}\, g/cm\, sec$ per liquidi come l’ acqua e di $1.8\, 10^{-4}\, g/cm\, sec$ per gas come l’aria. Mentre la viscosità statica vale $10^{-2}\, cm^2/sec$ per l‘ acqua, $0.15\, cm^2/sec$ per l’ aria e $1.2\, 10^{-3}$ per il mercurio.

Wiedemann- Franz , legge di: Il rapporto fra la conducibilità elettrica e quella termica dovuta ai soli portatori di elettricità è proporzionale alla temperatura assoluta e al rapporto $(R/F)^2$, ove R è la costante dei gas e F la costante di Faraday. Il fattore di proporzioonalità è circa 3. Questa legge ha avuto il ruolo di ”dimostrazione” della esistenza di elettroni liberi nei metalli responsabili della conducibilità elettrica, perchè sula base di questa ipotesi la cruda teoria di Drude, basata sulla trattazione degli elettroni con la statistica classica, prevedeva appunto la predetta relazione fra le conducibilità. L’ accordo si deve però considerare fortuito: e tuttavia la legge di W.-F. rimane valida anche se viene impiegata la statistica quantica (di Fermi Dirac) per gli elettroni (Sommerfeld): v. Conducibilità Elettrica.

Widom–Rowlinson, modello di: modello di gas a due componenti consistenti di particelle di tipo A e tipo B; si suppone che non vi sia interazione fra coppie di particelle A o fra coppie di particelle B; ma le particelle di tipo A interagiscono con quelle di tipo B attraverso un potenziale di cuore duro. Questo modello presenta una simmetria evidente per lo scambio del tipo. Se si considera uno stato di Gibbs in cui entrambe le specie hanno uguale potenziale chimico (quindi è conveniente pensare di usare l‘ insieme gran canonico per descrivere l‘ equilibrio), e se il potenziale chimico è abbastanza grande, questo modello presenta una transizione di fase in cui la simmetria AB è spontaneamente rotta e si può trovare o in una fase pura in cui la densità di A è alta, ovvero in cui è bassa (e quella di B è alta). E‘ stato il primo modello di gas non su reticolo per il quale è stato possibile dimostrare rigorosamante l‘ esistenza di una transizione di fase (Ruelle, 1971).

Wien , equazione di: in un corpo nero (v.) in equilibrio termico alla temperatura T la densità di energia della radiazione con frequenza fra ν e $\nu + d\nu$ è: $u_\nu = \nu^3 f(\nu/T)$ ove f è una opportuna funzione. Questa legge è conseguenza delle leggi dell' elettromagnetismo e della termodinamica. La determinazione della funzione f, invece, esula dalla Fisica Classica (v. Legge di Planck, Costante di Stefan–Boltzmann, MSC, Statistiche Quantiche).